Naval Research Laboratory
Washington, DC 20375-5000

NRL Report 9280

A Framework for the Detection of a Target of Unknown Velocity

M. J. STEINER

Target Characteristics Branch
Radar Division

September 27, 1990

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 27, 1990	Interim

4. TITLE AND SUBTITLE
A Framework for the Detection of a Target of Unknown Velocity

5. FUNDING NUMBERS
PE - PE61153N
TA - TARR021-05-43
WU - DN480-006

6. AUTHOR(S)
Steiner, Michael

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Research Laboratory
Washingon, DC 20375-5000

8. PERFORMING ORGANIZATION REPORT NUMBER
NRL Report 9280

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
Arlington, VA 22217

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The detection of a target in correlated Gaussian clutter, thermal noise, and extraneous interference is considered. The amplitude, phase, and Doppler frequency of the signal are not known a priori. The optimal detector, in the Neyman-Pearson sense is difficult to implement, and suboptimal methods are used. A general criterion that measures the performance of suboptimal detectors relative to an optimal test is presented. This criterion is encompassed in a design procedure used to design Doppler filters. The procedure allows many design considerations to be taken into account, which results in a design that attempts to minimize the number of filters required. For low dimensionality the procedure is of single filter designs; for higher dimensionality multiple filters are designed. The performance of these is compared with the results obtained by Emerson, Andrews, and the generalized likelihood ratio test. The clutter to thermal noise ratio and spectral width of the clutter are assumed known. When these parameters are not known exactly, the robustness of these and other filter systems is examined. A robustness theorem is presented, which relates the performance of a fixed filter to the clutter to thermal noise ratio as a function of the Doppler frequency. Future research is discussed.

14. SUBJECT TERMS
Detection — Filter design
Space-based radar

15. NUMBER OF PAGES 38

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

NSN 7540-01-280-5500 — Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std Z39-18 298-102

CONTENTS

A FRAMEWORK FOR THE DETECTION OF A TARGET OF UNKNOWN VELOCITY

1. INTRODUCTION

A basic problem in detection theory is that of detecting a target of unknown velocity. In each resolution cell of a pulsed radar, a detector is applied to determine whether interference only is present (H_0) or a (target) signal is also present (H_1). The interference is composed of the sum of thermal noise and clutter. Generally three signal parameters are not known a priori: amplitude, phase, and Doppler frequency. When the distribution of these variables is known a priori, the resulting hypothesis test can be solved through the Neyman-Pearson test. In this case, a likelihood ratio is averaged over the random parameters [1,2]. Although the resulting test satisfies the theoretical requirements of optimality, its complexity renders it too difficult to implement. Suboptimal methods are employed to approximate the performance of the average-likelihood ratio.

The generalized likelihood ratio test (GLRT) combines maximum likelihood estimates of the unknown parameters with the likelihood ratio [3]. This is discussed in Section 2, Eq. (6) for the radar problem. Essentially Eq. (6) requires finding the maximum over frequency of the magnitudes of the outputs of linear filters. Such a test is also difficult to implement because the maximization is over all possible values of frequency. Suppose the frequencies over which the maximization takes place are partitioned into a finite discrete set $\{f_1, f_2, \ldots, f_P\}$, the resulting test (finite constrained GLRT or CGLRT) now requires evaluation of only P linear filters. If P is low, the implementation requires substantially less hardware compared to the unconstrained test or GLRT. However, such a test may fail to perform close to what can optimally be achieved. Under such a constraint on the number of filters, it is necessary to find alternative means for designing filters that perform close to what can optimally be achieved. When the number of retuns N is small, a single filter may suffice to approximate the performance of the cumbersome average-likelihood test. These single filters are commonly referred to as moving target indicators (MTIs). Heuristically, the clutter is notched around zero Doppler, while moving targets and thermal noise are passed. For larger N, single filters are not as effective as multiple filters that can also integrate the target out of the thermal noise.

The design of single nonrecursive filters for low N is presently based on a technique first developed by Emerson and later expanded by Capon [4,5]. This procedure minimizes the interference output power or equivalently, maximizes the average signal to interference (SIR) ratio at the output of the filter when the target Doppler shift is uniformly distributed over all values [6]. The interference consists of the sum of thermal noise and clutter and typically is dominated by the clutter. The eigenvector corresponding to the minimum eigenvalue of the covariance matrix of the interference is the weight to be used in the filter. The interference output energy of such a (normalized) filter is the minimum eigenvalue of the covariance matrix of the interference. Andrews [7] has solved the design for multiple filters based on this criterion.

Although Emerson's criterion works well for many typical-interference power spectra, it fails when the power spectra are modified in other applications. The problems take the form of holes in the

Manuscript approved July 2, 1990.

Doppler coverage whereby certain target speeds cannot be detected. This is illustrated in Appendix A. Additionally, Emerson's procedure is unable to provide a desired target velocity response weighting. This latter problem has been addressed by modifying Emerson's filter by quadratic programming [8].

The limitations seen in Appendix A are inherent to the form of the criteria as discussed in Section 2. The purpose of this report is to present a criterion that does not have these limitations and produces filters with desirable characteristics. This criterion is presented in Section 3. In Section 4 the criterion is applied to the radar problem, and an equation is derived for the design of filters. This result is used in Section 5 to design single filters for small N. Comparisons are made with Emerson's method. In Section 6, multiple filters are constructed and the results are compared with Andrews' method and the CGLRT. In Section 7, the robustness of filters is examined relative to the clutter to thermal noise ratio C/N_0 and the normalized spectral width of the clutter. The results are summarized in Section 8.

2. DEFINITIONS

The classical hypothesis testing problem encountered when trying to discriminate a moving target from interference is represented as

$$H_0 : \mathbf{Y} = \mathbf{X}$$

$$H_1 : \mathbf{Y} = \mathbf{X} + se^{j\phi}\mathbf{S} \quad s > 0 \qquad \mathbf{S} = \begin{pmatrix} 1 \\ e^{j\theta} \\ \vdots \\ e^{j(N-1)\theta} \end{pmatrix}, \tag{1}$$

where N denotes the dimension. The signal vector is $\mathbf{S}$. The vector $\mathbf{X}$ consists of the sum of thermal noise and clutter and is assumed to have zero mean and a Gaussian distribution. A covariance matrix R associated with $\mathbf{X}$ is defined as $R_{j,k} = E(x_j^* x_k)$, where * denotes complex conjugate. The diagonal elements of R are fixed at σ^2, which is the sum of the thermal and clutter energy per component. The unknown constant phase ϕ of the received signal is assumed to have a uniform distribution in the interval $[0, 2\pi]$. The Doppler frequency θ is also assumed random with a uniform distribution in the interval $[0, 2\pi]$. The signal amplitude s is assumed to be greater than zero; the prior distribution of s is unknown. The probability of $\mathbf{Y}$ given s, ϕ, θ assuming the thermal noise is Gaussian distributed is

$$P(\mathbf{Y} \mid s, \phi, \theta) = \frac{1}{(\pi)^N \mid R \mid} e^{-(\mathbf{Y} - se^{j\phi}\mathbf{S})^T R^{-1} (\mathbf{Y} - se^{j\phi}\mathbf{S})^*}, \tag{2}$$

where T denotes vector transpose. The likehood ratio is,

$$L(\mathbf{Y} \mid s, \phi, \theta) = \frac{P_1(\mathbf{Y} \mid s, \phi, \theta)}{P_0(\mathbf{Y})}. \tag{3}$$

The optimal test as discussed in Section 1 is the Neyman-Pearson or average-likelihood ratio test [9] that compares the average-likelihood with a threshold. The average likelihood can be written as

$$L(\mathbf{Y} \mid s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} L(\mathbf{Y} \mid s, \phi, \theta) P(\phi)\, P(\theta) d\phi d\theta$$

$$= \int_{0}^{2\pi} e^{-s^2 \mathbf{S}^T R^{-1} \mathbf{S}^*} I_0(2s \mid \mathbf{Y}^T R^{-1} \mathbf{S}^* \mid) \frac{1}{2\pi} d\theta. \tag{4}$$

This detector is difficult to implement. The test can be approximated and considerably simplified by using the GLRT that compares

$$\max_{s,\phi,\theta} L(\mathbf{Y} \mid s, \phi, \theta) \tag{5}$$

to a threshold. This can be simplified for our problem to

$$\delta(\mathbf{Y}) = \begin{cases} 1 & \qquad > \\ 0 \text{ or } 1 \quad \text{if } \max_\theta K(\theta) \mid \mathbf{Y}^T R^{-1} \mathbf{S}^* \mid & \qquad = \tau \\ 0 & \qquad < \end{cases} \tag{6}$$

where $K(\theta) = 1/\sqrt{\mathbf{S}^T R^{-1} \mathbf{S}^*}$. This test has a number of desirable properties when the frequency θ is *fixed*, (1) it can be accomplished by taking the magnitude of a linear filter, (2) it is equivalent to the average likelihood test in Eq. (4) hence is optimal for fixed frequency, and (3) it possesses the desirable property of being uniformly most powerful (UMP) with respect to the unknown amplitudes. Although methods exist for finding the maximum of the above equation over θ, they are difficult to implement. Estimation methods for finding θ for a similar problem are given in Ref. 10. However, if $\mid \mathbf{Y}^T R^{-1} \mathbf{S}^* \mid$ is implemented for each θ, a matched filter needs to be built or evaluated digitally for each θ. This requires an infinite number of filters. Therefore the test is typically approximated by a finite number of filters. In this case each filter must cover an interval of frequencies θ. Degradation, relative to the infinite bank of filters, Eq. (6), will occur at frequencies where the finite filters become mismatched. An optimality or design criterion to measure this degradation can be defined, and new coefficients can be determined. The amount of the deterioration with frequency can vary greatly, depending on the filters used, the covariance matrix, and on the length N. The best filter design when the number of filters is limited along with the associated design criteria is paramount to performance the cost.

An optimal criterion that results from hypothesis testing is the maximum of the average probability of detection (P_d) over the frequency range. For independent and identically distributed (IID) Gaussian interference with a Raleigh distribution on the amplitudes, Brennan et al. [1] show that the average likelihood ratio test (which maximizes the average probability of detection) and the generalized likelihood ratio test perform in nearly the same manner. Little has been done to design single filters when the criterion of maximizing the average probability of detection is desired when correlated interference is present. In fact this is not a good criterion. Since the problem of correlated Gaussian interference, P_d may range over several orders of magnitude depending on the frequency of the target. Hence an average of these values over the frequency range is strongly dominated by the largest values. The frequencies where the largest values occur may be over a narrow interval that may not even be considered important. The second criterion that predominates in the literature is the one of Emerson, which consists of minimizing the average interference power at the output of the filter. The problems associated with this criterion are discussed in Appendix A.

An optimal criterion is needed that can be used to design either a single filter or a bank of filters that do not suffer the limitations of the previous methods. The properties that we consider necessary for the criterion to satisfy are the following:

- The criterion should accurately reflect the degradation of the desired filters relative to what can be achieved in an optimal sense;
- The criterion should be broad enough to allow parameters of the design (such as number of filters, positioning of the filters, etc.) to be varied;
- The criterion should be capable of being weighted to take a particular desired profile or cost function into consideration.

For the criterion to reflect the degradation of the bank of filters, a suitable distance measure must be defined. An appealing indicator at a particular frequency θ is P_d for a given probability of false alarm (P_{fa}). Obviously we would like P_d to be as close to what can be optimally achieved at a given a P_{fa}. To determine what values are actually close, an energy function can be introduced. This function should tell us how much extra energy is needed for a suboptimal processor to perform as well as an optimal processer as a function of frequency. Recently, Compernolle [11] has brought to attention a distance measure for use in speech recognition for measuring the difference between power spectrums. This measure is based upon the difference of the two power spectra in the log domain. Similarly, the radar engineer is interested in the signal energy and noise energy expressed in decibels. The reason for preferring decibels over absolute energy values relates to the radar range equation and the exponential nature of the probability of error equations. A more relevant distance measure is based on the log energy. In effect, it is desirable to know as a function of θ the necessary signal energy in dB required of the suboptimal processor to perform as well as the optimal processor. A criterion based on this measure is introduced and used in the remainder of the report.

3. MINIMAX LOG ENERGY CRITERIA (MMLE)

A general optimal criterion that satisfies the properties we required of an optimal criterion is now presented in Section 2. The input SIR s^2/σ^2 is denoted as γ.

Suppose, for each frequency $\theta \in [0, 2\pi]$ the Neyman-Pearson design criteria is used to find the most powerful tests $\delta_\theta^\gamma(\mathbf{Y})$, where γ denotes the input SIR, under the constraint $P_{fa}(\delta_\theta^\gamma(\mathbf{Y})) = \alpha$. Then any rule $\delta_c^\gamma(\mathbf{Y})$ will, by the Neyman-Pearson lemma, satisfy

$$P_d(\delta_c^\gamma(\mathbf{Y}) \mid \Theta = \theta) \leq P_d(\delta_\theta \gamma(\mathbf{Y}) \mid \Theta = \theta); \tag{7}$$

whenever $P_{fa}(\delta_c^\gamma(\mathbf{Y}) \mid \Theta = \theta) \leq \alpha$, Θ denotes the actual frequency of the signal.* The rule $\delta_c^\gamma(\mathbf{Y})$ is assumed to be of the form

$$\delta(\mathbf{Y}) = \begin{cases} 1 & > \\ 0 \text{ or } 1 \text{ if } \max_i |\mathbf{C}_i^T \mathbf{Y}| & = \ \tau, \\ 0 & < \end{cases} \tag{8}$$

*We assume P_d, P_{fa} can be written in terms of γ to not depend on s or σ separately.

where $i = 1, \ldots, r$, r being the number of filters. Each filter coefficient vector $\mathbf{C}_i$ will be associated with a frequency range of interest $\Lambda_i \subset [0, 2\pi]$ such that $\{\Lambda_1, \ldots, \Lambda_r\}$ forms a partition of the interval $[0, 2\pi]$. Suppose a filter with coefficients $\mathbf{C}$ is to be designed for and interval $\Lambda_A \in \{\Lambda_1, \ldots, \Lambda_r\}$. It is desired that $P_d(\delta_c^{\gamma}(\mathbf{Y}) \mid \Theta = \theta)$ be close in some sense to the optimum $P_d(\delta_\theta^{\gamma}(\mathbf{Y}) \mid \Theta = \theta)$ for each $\theta \in \Lambda_A$.

Definition 1 — *Define γ_c as the input SIR required such that $\delta_c^{\gamma_c}(\mathbf{Y})$ will have the same P_0 as $\delta_\theta^{\gamma}(\mathbf{Y})$ at $\Theta = \theta$, i.e.,*

$$P_d(\delta_c^{\gamma_c}(\mathbf{Y}) \mid \Theta = \theta) = P_d(\delta_\theta^{\gamma}(\mathbf{Y}) \mid \Theta = \theta),$$

when $P_{fa} = \alpha$ for *both tests. The* log *energy $e(\theta)$ is given as*

$$e(\theta) = \log \gamma_c - \log \gamma.$$

Furthermore the log energy risk is defined as

$$r(\delta) = \max_{\theta \in \Lambda_A} g(\theta)\, e(\theta),$$

where $g(\theta)$ is an arbitrary weighting function.

The log energy function $e(\theta)$ is the log energy required for the filter with coefficients $\mathbf{C}$ to perform the same as the optimal filter at the given frequency θ. The log energy risk is defined as the minimum log energy needed for the suboptimal detector to perform at least as well as the δ_θ^{γ} $(\mathbf{Y})$ for all $\theta \in \Lambda_A$. This risk can be weighted depending on the cost structure desired on the set Λ_A. The weighting will take the form of a function $g(\theta)$. The risk can be minimized in a Bayesian sense,

$$\delta \in \arg \min_{\delta} \max_{\theta \in \Lambda_A} g(\theta) e(\theta)$$

or this can be written in terms of the filter $\mathbf{C}$,

$$\mathbf{C} \in \arg \min_{\mathbf{C}} \max_{\theta \in \Lambda_A} g(\theta) e(\theta). \tag{9}$$

The log energy $e(\theta)$ accurately reflects the degradation of the desired filters relative to what can be achieved in the optimal case in a manner relevant to radar engineering. The number of filters that can be designed by using this criterion is not constrained, nor is the frequency separation of the filters, as the intervals Λ can arbitrarily be chosen. Thus, the criterion is broad enough to allow important design parameters to be varied. Finally, the criterion is capable of being weighted by the arbitrary function $g(\theta)$, which allows it to incorporate desired cost functions such as velocity profiles.

It is interesting to compare this criterion with the maximum formulation used in many robustness problems [12,13] in which one seeks a filter to maximize the minimum gain over a class of possible signals and/or interference models,

$$\max_{\mathbf{C}} \min_{\theta \in \Lambda_A} \mathrm{SIR}. \tag{10}$$

This theory has been developed so that explicit solutions are available under certain hypothesis. In our case the signal may vary by taking on different θ. This criterion is compared with the MMLE in the Section 5 for the design of single filters.

4. MMLE APPLIED TO THE RADAR PROBLEM

To apply the MMLE to the radar problem Eq. (1), $e(\theta)$ must be determined in Eq. (9). This requires the rules $\delta_\theta^\gamma(\mathbf{Y})$ and the associated P_d and P_{fa}. The most powerfull tests δ_θ^γ by the Neyman-Pearson criterion are found to be a form of the Wiener filter.

$$\delta(\mathbf{Y}) = \begin{cases} 1 & > \\ 0 \text{ or } 1 \text{ if } |\mathbf{Y}^T R^{-1}\mathbf{S}^*| & = \quad \tau(\theta) . \\ 0 & < \end{cases} \tag{11}$$

The P_d for this test is Ref. 14

$$Q\left(\frac{l}{r}, \sqrt{2 \ln \frac{1}{P_{fa}}}\right),$$

where $l = |\mathbf{C}^T\mathbf{S}|$, $r^2 = \mathbf{C}^\dagger R\mathbf{C}$, Q is the Marcum Q function. We wish to find the input SIR γ_c such that the suboptimal rule $\delta_c^{\gamma_c}(\mathbf{Y})$ performs the same as the test $\delta_\theta^\gamma(\mathbf{Y})$ for some $\theta \in \Lambda_A$. This requires,

$$Q\left(\frac{l(\theta)}{\tau(\theta)}, \sqrt{2 \ln \frac{1}{P_{fa}}}\right) = Q\left(\frac{l_c}{\tau_c}, \sqrt{2 \ln \frac{1}{P_{fa}}}\right) \quad \theta \in \Lambda_A ,$$

where l_c, τ_c correspond to the test $\gamma_c^{\gamma_c}(\mathbf{Y})$ and $l(\theta)$, $\tau(\theta)$ to the respective $\delta_\theta^\gamma(\mathbf{Y})$. This amounts to

$$\frac{l(\theta)}{e(\theta)} = \frac{l_c}{\tau_c} \quad \theta \in \Lambda_A ,$$

which is equivalent to

$$G_\theta \gamma = G_c \gamma_c ,$$

where

$$G(\theta) = \frac{1}{\gamma} \frac{|\mathbf{C}^T\mathbf{S}|^2}{\mathbf{C}^\dagger R\mathbf{C}} \tag{12}$$

is the gain for a filter with coefficients $\mathbf{C}$, and $\dagger$ denotes conjugate transpose. G_θ, G_c denotes the gains for the most powerful test and the suboptimal test respectively. Now $e(\theta)$ can be determined from

$$e(\theta) = \log \gamma_c - \log \gamma = \log G_\theta - \log G_c ,$$

and substituting $R^{-1}\mathbf{S}^*$ for the coefficients of the optimal test $\delta_\theta^\gamma(\mathbf{Y})$ yields

$$e(\theta) = \log \mathbf{S}^T R^{-1}\mathbf{S}^* \frac{1}{\gamma} - \log \frac{|\mathbf{C}^T\mathbf{S}|^2}{\mathbf{C}^\dagger R\mathbf{C}} \frac{1}{\gamma}$$

$$= \log \mathbf{S}^T R^{-1}\mathbf{S}^* \frac{\mathbf{C}^\dagger R\mathbf{C}}{|\mathbf{C}^T\mathbf{S}|^2} . \tag{13}$$

To find the log energy risk, $e(\theta)$ is maximized with respect to θ. Assuming uniform costs, i.e., $g(\theta) = 1 \ \forall \theta \in \Lambda_A$,

$$r(\delta) = \max_{\theta \in \Lambda_A} e(\theta)$$

$$= \mathbf{C}^\dagger R \mathbf{C} \max_{\theta \in \Lambda_A} \frac{\mathbf{S}^T R^{-1} \mathbf{S}^*}{|\mathbf{C}^T \mathbf{S}|^2}. \tag{14}$$

To find the rule or suboptimal coefficient vector $\mathbf{C}$, the log energy risk is minimized.

$$\min_{\mathbf{C}} \left[\mathbf{C}^\dagger R \mathbf{C} \max_{\theta \in \Lambda_A} \frac{\mathbf{S}^T R^{-1} \mathbf{S}^*}{|\mathbf{C}^T \mathbf{S}|^2} \right]. \tag{15}$$

A filter with coefficients $\mathbf{C}$ is obtained when the minimum is reached (it may or may not be unique). We call filters derived in such a manner MMLE filters. Note that the maximum over θ is essentially the maximum of the ratio of two polynomials with coefficients from $\mathbf{C}$ or R. No general solution appears other than a search algorithm. Numerical algorithms have been developed and are discussed in Appendix C.

5. DESIGN FOR LOW-DIMENSION N

The results shown in the preceding section will now be used to design a detector incorporating a single filter, when the dimension N is low. The degree to which N is considered low is relative to the particular interference covariance matrix under consideration, hence may change for different applications.

We will consider three scenarios and the corresponding covariance matrices.

(a) The received interference is dominated by clutter introduced by rain, cloud, foliage, and the like. The clutter is highly correlated and will be modeled as a Gaussian process with a Gaussian-shaped clutter power spectrum. The covariance matrix can be written as in Ref. 15,

$$R = \begin{pmatrix} 1 + \sigma_X^2 & \rho(T) & \cdots & \rho(NT) \\ \rho(T) & 1 + \sigma_X^2 & \cdots & \vdots \\ \vdots & \vdots & & \vdots \\ \rho(NT) & \rho((N-1)T) & \cdots & 1 + \sigma_X^2 \end{pmatrix}$$

where

$$\rho(kT) = e^{-2(\pi k \sigma_C T)^2}. \tag{17}$$

The thermal noise energy per dimension is σ_X^2, the clutter energy per dimension is normalized to 1, and the pulse repetition is T. The total interference energy per dimension σ^2 is then $1 + \sigma_X^2$. The standard deviation of the Gaussian clutter power spectrum is σ_C, and the normalized spectral width of the spectrum is $\sigma_C T$.

(b) Extraneous interference in a radar return may cause one or more components to be corrupted. For simplicity, one component will be assumed corrupted. It is assumed that the interference dominates so that no useful information can be extracted from the corrupted pulses in the detection process. The optimal solution for a fixed frequency and ith pulse corrupted is to remove the ith row and ith column of the covariance matrix above and accordingly process the remaining pulses [16]. The operator $P_i(\mathbf{X})$ denotes the vector $\mathbf{X}$ with the ith component removed.

c. A covariance matrix is devised in Appendix A to illustrate the Emerson procedure. This matrix can be written as,

$$R = \sum_i \lambda_i \frac{\mathbf{S}_i \mathbf{S}_i^\dagger}{N} \quad \mathbf{S}_i = \begin{bmatrix} 1 \\ e^{j\frac{2\pi(i-1)}{N}} \\ e^{j\frac{4\pi(i-1)}{N}} \\ \vdots \\ e^{j\frac{2\pi(N-1)(i-1)}{N}} \end{bmatrix}, \tag{18}$$

where $i = 1 \ldots, N$. The set $\{\lambda_1, \ldots, \lambda_N\}$ contains the eigenvalues of the matrix R, and the eigenvalues are assumed distinct. Note that the matrix is Toeplitz.

For each of the three cases, the Emerson and the MMLE methods are compared with the GLRT or infinite filter bank, Eq. (6). The gain function for the GLRT is optimal for a given θ. Henceforth it is referred to as the optimal gain. The GLRT as a detector however is not optimal when one does not have a priori knowledge of θ. We assume that the speed of the target is unknown and is uniformly distributed in the interval $[0, 2\pi]$. Furthermore we assume a uniform cost function $g(\theta)$ in Eq. (9). For simplicity, $\sigma_C T$ is fixed at 0.05 and C/N_0 at 50 dB. The design goal is to be as close to the optimal gain in the passband as possible. Considering $\sigma_C T = 0.05$, a somewhat arbitrary passband will be assumed between the normalized frequency ($f = \theta/2\pi$) intervals $f \in [.1, .9]$ and $f \in [.2, .8]$ depending on the remaining parameters. If a single filter is not sufficient to cover the passband satisfactorily, it is necessary to either change the passband or add additional filters as discussed in the next section.

In Fig. 1, gain is plotted vs f, for the covariance matrix, Eq. (16) for $N = 4$. The optimal gain obtained by using the filter coefficients $R^{-1}\mathbf{S}^*$ is plotted along with the single filter gains of Eq. (12) based on Emerson's weights and the MMLE. It is clear, in the passband, that both single filters perform nearly the same as the GLRT, Eq. (6). Nearly the same performance is observed for $N = 6$ shown in Fig. 2, except the MMLE begins to separate from Emerson's filter at the beginning and middle of the passband. This separation can be seen clearly for $N = 8$ in Fig. 3 where the MMLE achieves at most 20 dB more improvement over Emerson's method for frequencies in the intervals $f \in [.15, .35]$, $f \in [.65, .85]$, and sacrifices at most 5 dB in the interval $f \in [.35, .65]$. The maximin formulation of Eq. (10) is shown in Fig. 4 relative to the MMLE and infinite bank Eq. (6) for $N = 8$. Both the MMLE and maximin were optimized over $f \in [.2, .8]$. The relatively poor performance of the maximin is heuristically due to maximizing the minimum gain that occurs at $f = .2$ and $f = .8$ in the interval $f \in [.2, .8]$. This is done at the expense of the remaining passband that is allowed to sharply degrade. A more relevant application of the maximin procedure is discussed in Section 8.

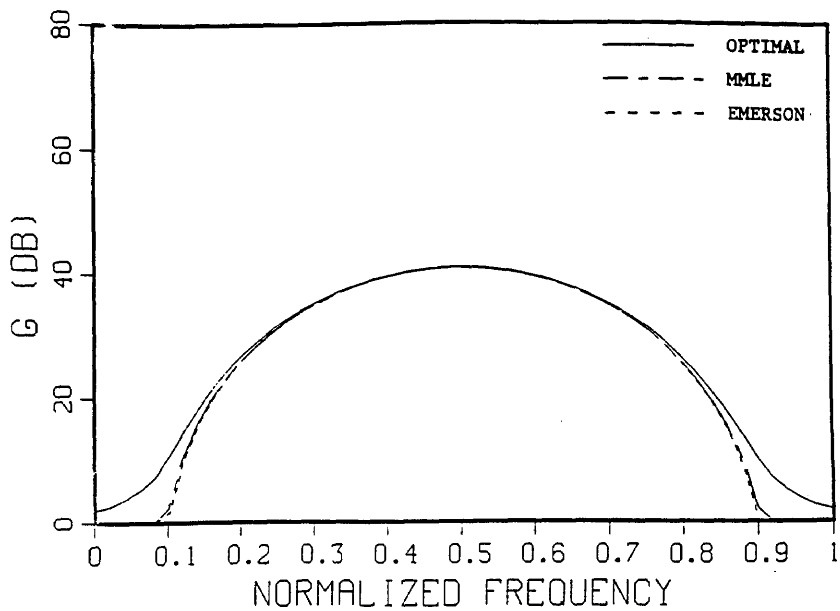

Fig. 1 — Gain-Gaussian clutter power spectrum $N = 4$, $\sigma_C T = 0.05$, $C/N_0 = 50$ dB

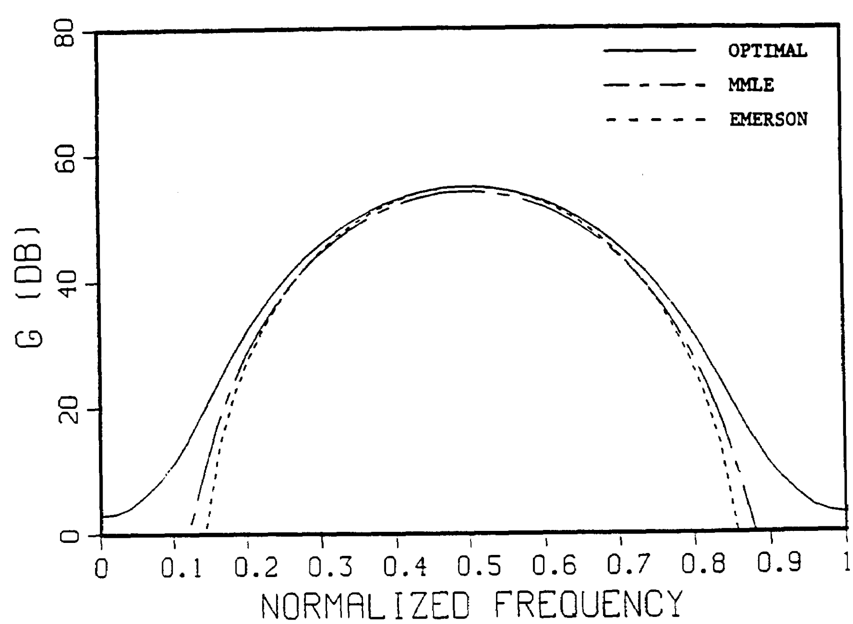

Fig. 2 — Gain-Gaussian clutter power spectrum $N = 6$, $\sigma_C T = 0.05$, $C/N_0 = 50$ dB

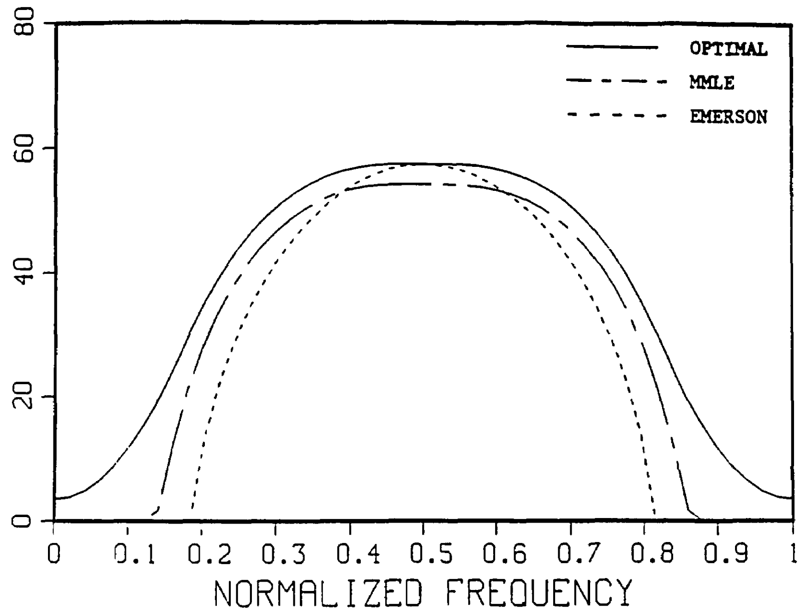

Fig. 3 — Gain-Gaussian clutter power spectrum $N = 8$, $\sigma_C T = 0.05$, $C/N_0 = 50$ dB

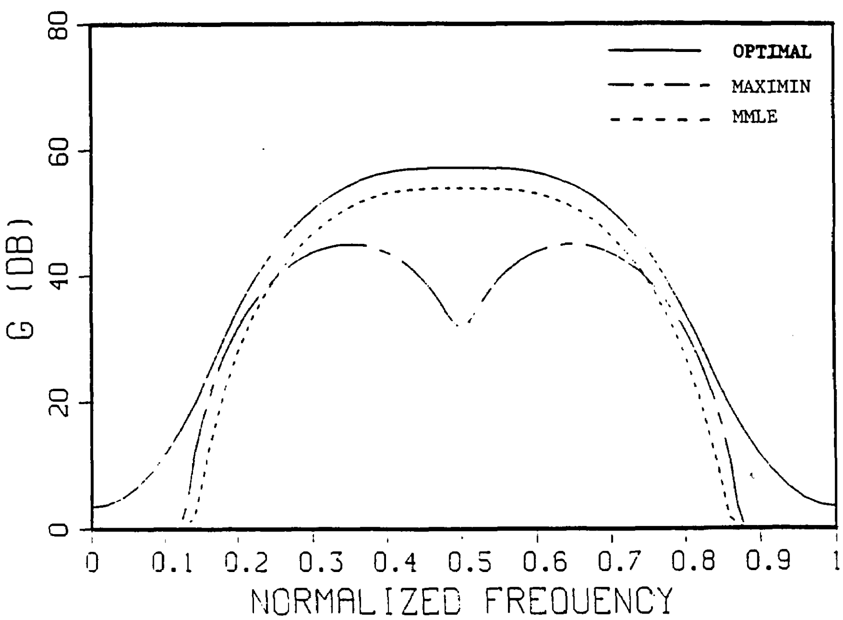

Fig. 4 — Gain-Gaussian clutter power spectrum $N = 8$, $\sigma_C T = 0.05$, $C/N_0 = 50$ dB

In Fig. 5, curves of the optimal gain, MMLE gain, and gain obtained by using Emerson's filter are shown for $N = 5$. The third row and column of Eq. (16) are removed to account for extraneous interference. The gain function for Emerson's filter yields a blind speed at $f = .5$, which could be critical for many systems (this problem has been observed for odd N). The MMLE maintains a nearly identical loss across the passband. The single MMLE filter again yields uniform coverage in Fig. 6 for $N = 9$ while Emerson's filter produces poor results at $f = 0.5$.

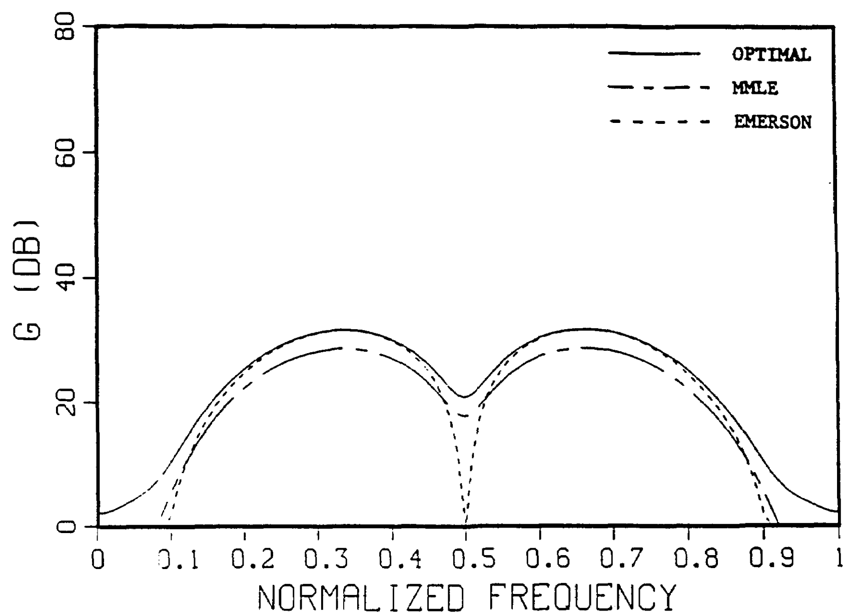

Fig. 5 — Gain-extraneous interference P_3, $N = 5$, $\sigma_C T = 0.05$, $C/N_0 = 50$ dB

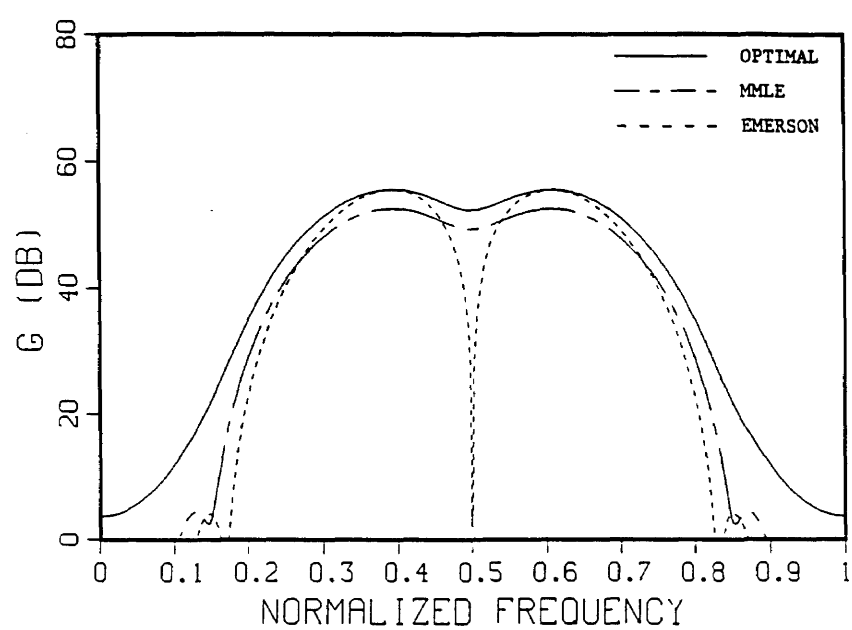

Fig. 6 — Gain-extraneous interference P_5, $N = 9$, $\sigma_C T = 0.05$, $C/N_0 = 50$ dB

We now consider the covariance matrix of the form of Eq. (18). Again Emerson's method and the MMLE are compared with the infinite bank. For $N = 3$ shown in Fig. 7, the eigenvalues are chosen so that most of the interference is near zero Doppler; $\lambda_1 = 1.$, $\lambda_2 = 0.01$, and $\lambda_3 = 0.0099$. As explained in Appendix A, choosing any eigenvector of the covariance matrix results in two nulls in the Doppler space, in our case at $f = 0$, $f = 2/3$. The gain for the MMLE design is uniformly close to the optimal gain. The same effect is seen in Fig. 8 in which $N = 5$. Here $\lambda_1 = 1.$, $\lambda_2 = 0.0000999$, $\lambda_3 = 0.0001$, $\lambda_4 = 0.0001$, $\lambda_5 = 0.0001$. These cases are based on a covariance matrix that is designed to help in understanding the problem and may not be indicative of covariance matrices encountered in physical processes.

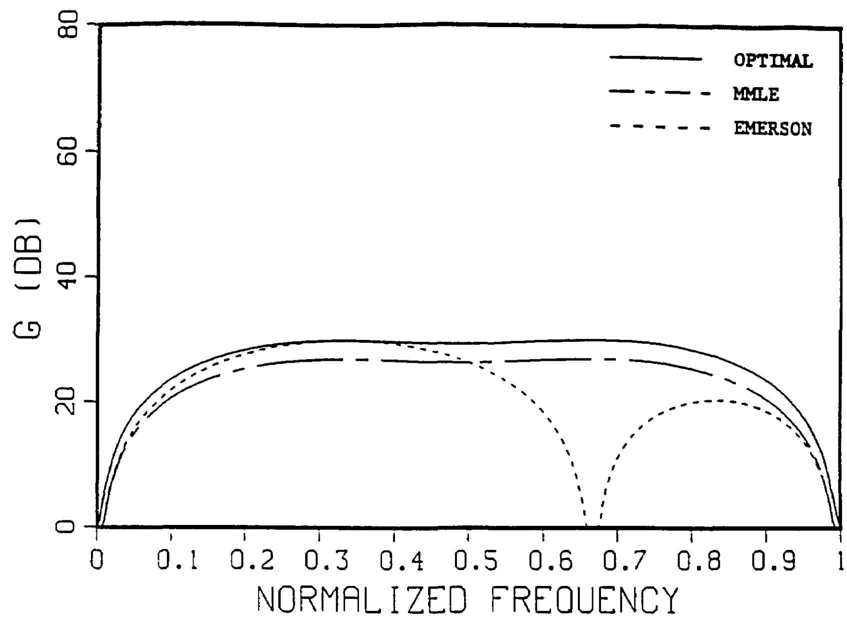

Fig. 7 — Gain-interference matrix (17), $N = 3$

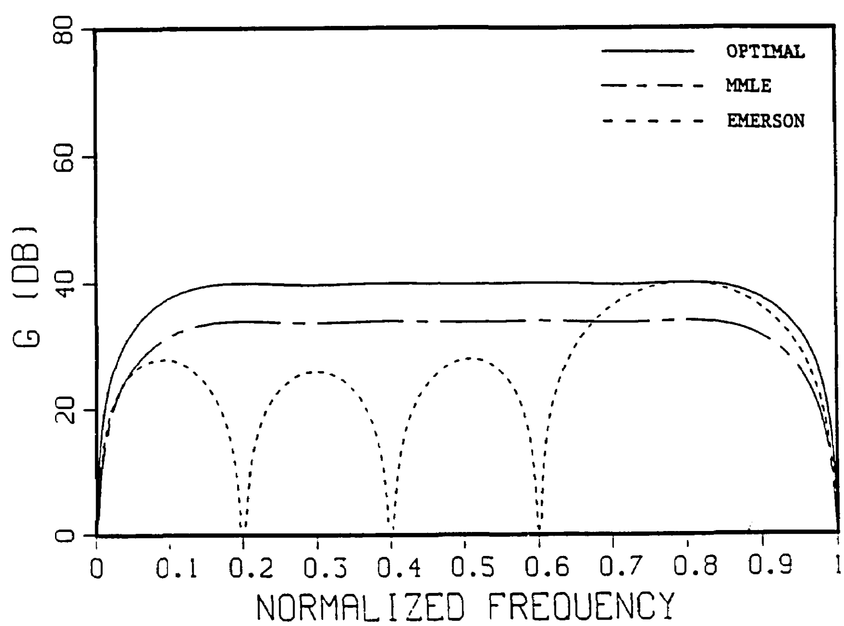

Fig. 8 — Gain-interference matrix (17), $N = 5$

6. DESIGN FOR HIGH-DIMENSION N

When the single filter designs of the previous section fail to achieve the desired goal, a useful recourse is to consider the implementation of multiple filters. To compare the designs, we define the coverage $C(f)$ as,

$$C(f) = \max_i G_i(2\pi f), \quad i = 1, \ldots, P \tag{19}$$

where $G_i(2\pi f)$ is the associated gain function, Eq. (12) of the ith filter, and P is the number of filters. The uniform weighting function, $g(\theta) = 1$ in the passband; $g(\theta) = 0$ otherwise, is assumed. It is desirable to have the coverage as close in the passband to what is obtainable with the infinite bank of filters, Eq. (6) where each filter is optimal for a corresponding frequency. The covariance matrices that are considered are cases 1 and 2 of Section 5, that of the interference having a Gaussian power spectrum and extraneous interference corrupting one component of the received vector. Again we assume $\sigma_C T = 0.05$ and $C/N_0 = 50$ dB.

Comparisons are made between the MMLE the N filters by Andrews, and the approximated CGLRT with N equally spaced filters. Andrews' design for multiple filters is a straightforward extension of the Emerson's procedure. Here the Doppler space is partitioned into N equally spaced intervals, and N filters are found by maximizing the average gain for each respective interval. Specifically the ith interval is given by

$$\left[\frac{(2i-1)}{2N}, \frac{(2i+1)}{2N}\right], \tag{20}$$

where $i = 0, \ldots, N-1$ and the units are normalized frequency. This is in contrast to Emerson's procedure that yields a single filter by maximizing the average gain over the entire [0, 1] normalized frequency interval. The CGLRT consists of N filters with coefficients of the form $R^{-1}\mathbf{S}^*$. By partitioning the Doppler space into N equally spaced intervals in Eq. (20) and by using the average frequency in each interval to evaluate the signal, N such filters are derived. The MMLE is well suited for the problem of multiple filter design, since filters are optimized over disjoint frequency intervals Λ_i as described in Section 3. To meet a design goal with the least number of filters, one must determine the frequency intervals Λ_i. Given the same general design goals and passbands as in Section 3, this can be done as follows. Assuming a single filter did not meet the design requirements, the first interval Λ_1 is set equal to the interval $f \in [0, .5]$. The MMLE is applied and the gain is determined. If the gain is not sufficiently close to the optimal gain in the interval $f \in [0, .5]$, the interval is decreased to $f \in [0, x]$ where x is less then 0.5, and the MMLE is applied again. Changes (both increasing and decreasing as necessary) are continued in this manner until the gain meets the design requirements in the interval. Suppose the single filter now meets the requirements in the interval $f \in [0, x]$, and a second filter is designed in the interval $f \in [x, 1]$ and the coverage, Eq. (19), of the first and second filters is compared with the design requirements. This process of changing the interval is continued as above ($f \in [x, y]$ y variable) until the coverage is satisfactory through the second interval. Enough filters are added by following this procedure until the passband is covered to meet the requirements.

Figure 9 shows the approximated CGLRT vs the GLRT or infinite filter bank, Eq. (6), for $N = 10$. The coverage provided in the passband is practically identical to the GLRT coverage. This coverage is repeated in Fig. 10 in which filters are designed by using the Andrews' method. Finally the MMLE is shown in Fig. 11 in which only four filters are used. The intervals determined by the method above are $f \in [.17, .4]$, $f \in [.4, .5]$; and conjugate filters are used for the intervals

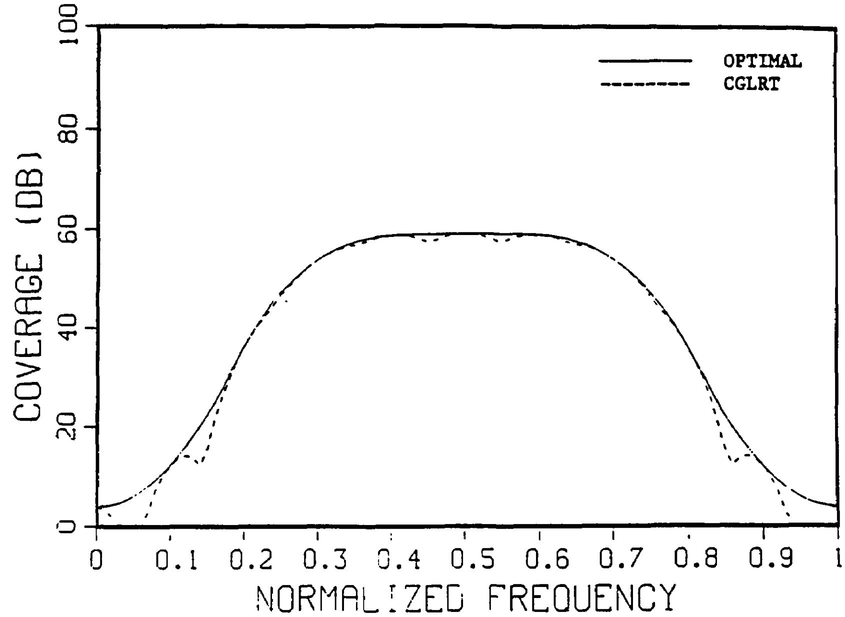

Fig. 9 — Coverage-Gaussian clutter power spectrum multiple filters CGLRT, $N = 10$, $\sigma_C T = 0.05$, $C/N_0 = 50$ dB

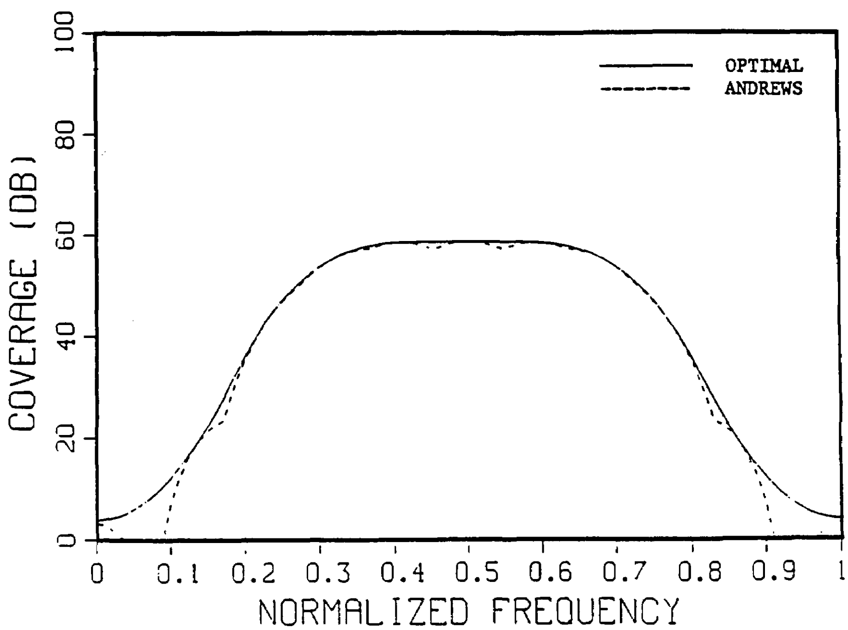

Fig. 10 — Coverage-Gaussian clutter power spectrum multiple filters Andrews, $N = 10$, $\sigma_C T = 0.05$, $C/N_0 = 50$ dB

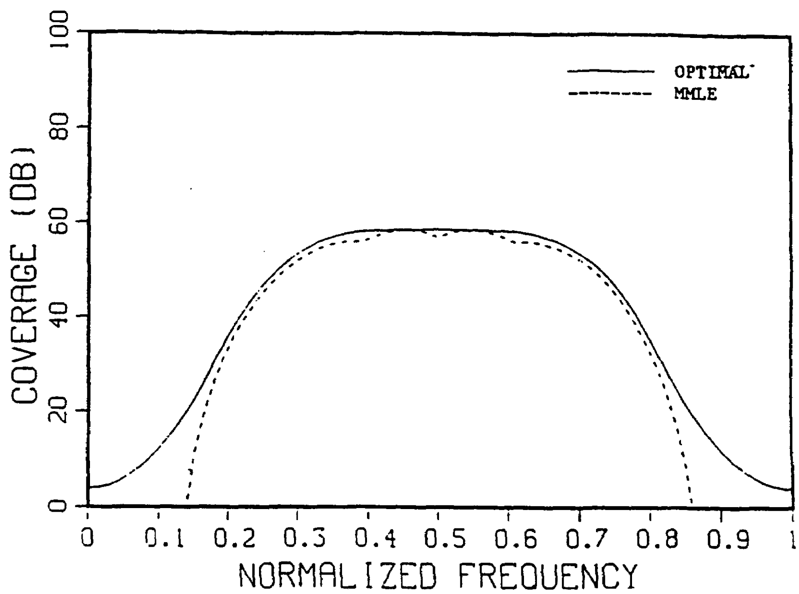

Fig. 11 — Coverage-Gaussian clutter power spectrum multiple filters MMLE, $N = 10$, $\sigma_C T = 0.05$, $C/N_0 = 50$ dB

$f \in [.5,.6]$, $f \in [.6,.83]$. Here the length of the interval of the first filter is 0.23 while the second is less then half of this at 0.1. Nonuniform lengths are important if a minimum number of filters is desired.

Extraneous interference is considered. Figures 12 and 13 compare the CGLRT and Andrews' filters respectively to the infinite bank. The dimension $N = 9$ and the middle component is corrupted, resulting in eight pulses. The coverage provided in the passband is nearly the same as the infinite filter bank coverage. Figure 14 shows the MMLE design in which only three filters were needed to approximate the optimal gain. The intervals used in the design were $f \in [.17, .38]$, its conjugate filter interval $f \in [.62, .83]$, and $f \in [.38, .62]$. The lengths of the intervals are nearly equal.

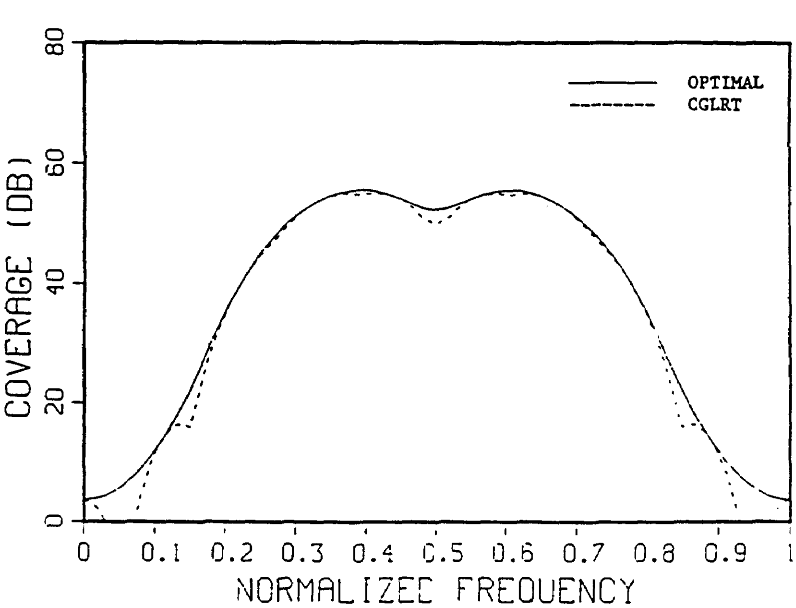

Fig. 12 — Coverage-extraneous interference P_5, multiple filters CGLRT, $N = 9$, $\sigma_C T = 0.05$, $C/N_0 = 50$ dB

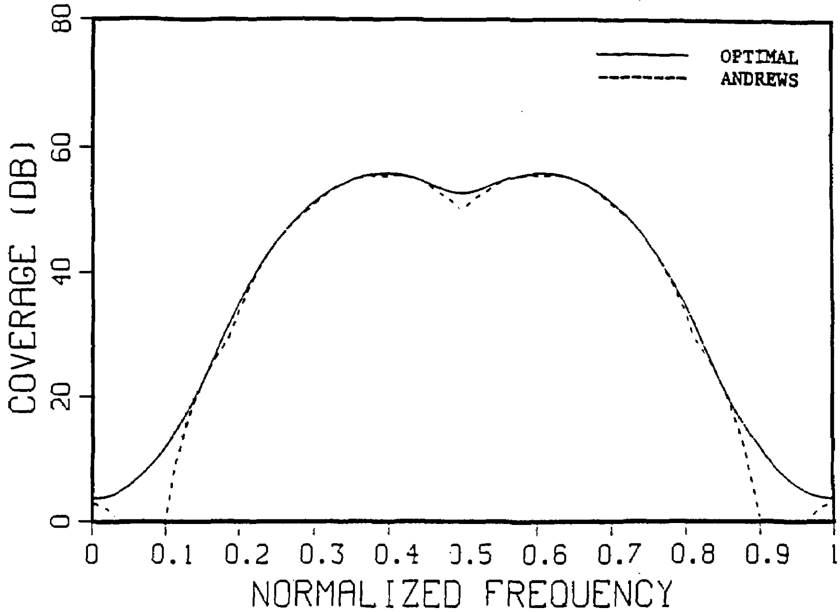

Fig. 13 — Coverage-extraneous interference P_5, multiple filters Andrews, $N = 9$, $\sigma_C T = 0.05$, $C/N_0 = 50$ dB

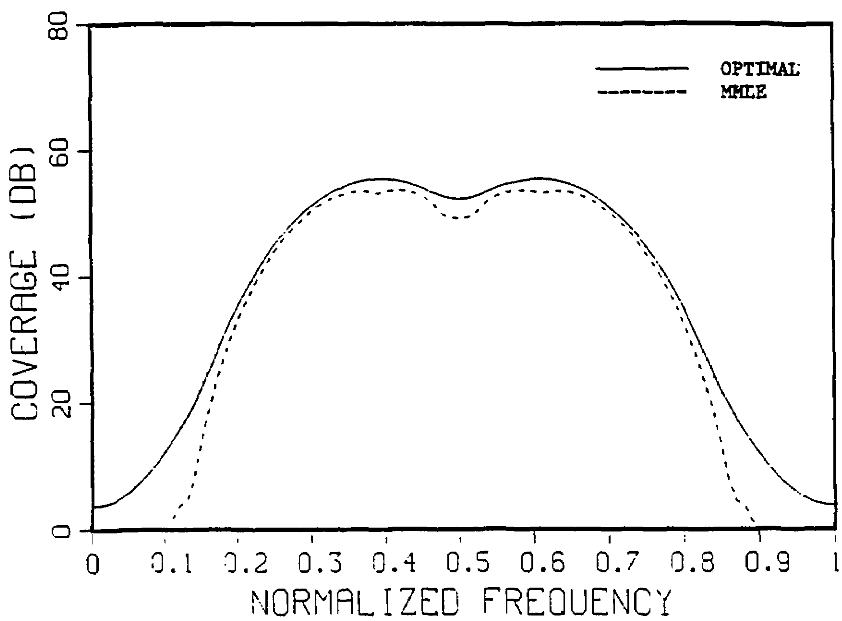

Fig. 14 — Coverage-extraneous interference P_5, multiple filters MMLE, $N = 9$, $\sigma_C T = 0.05$, $C/N_0 = 50$ dB

In summary we have shown that only four filters are needed to effectively cover the passband for $N = 10$. For extraneous interference only three filters are necessary.

7. ROBUSTNESS

The filters in the previous section were designed under the assumption that the statistics of the interference are accurately reflected by the interference covariance matrix used. This may be so in certain space based applications. However, for other applications this may not be so. In fact, the interference matrix may change from range cell to range cell and from scan to scan of a pulsed radar. The received signal may either possess Gaussian or nongaussian statistics. For our purposes a Gaussian process is assumed and furthermore we assume that the received signal, Eq. (1) (which is at a constant range), is wide sense stationary. The actual covariance matrix associated with the Gaussian interference is assumed to lie within a class of covariance matrices that have a Gaussian clutter power spectrum. The class $\mathscr{R}$ of covariance matrices is defined as

$$\mathscr{R} = \left\{ R \text{ as in Eq. (16)} \mid \sigma_C T \in \Delta_a, \frac{C}{N_0} \in \Delta_b \right\}, \tag{21}$$

where Δ_a is the set of relevant normalized spectral widths, and Δ_b is the set of relevant clutter to thermal noise ratios, $C/N_0 = 1/\sigma_X^2$. Relationships on filter performance when C/N_0 is varied are presented. The robustness of the MMLE, Andrews' method, and the CGLRT filters will be examined when they are designed at a fixed operating point of $\sigma_C T$ and C/N_0.

We begin with an equation derived from Ref. 7 that relates gain averaged over the entire normalized frequency interval [0, 1], termed the improvement factor IF, when Emerson's method is used,

$$\frac{1}{I_{cn}} = \left[\frac{\frac{C}{N_0}}{1 + \frac{C}{N_0}} \right] \frac{1}{I_c} + \frac{1}{1 + \frac{C}{N_0}}. \tag{22}$$

I_c represents the IF when the eigenvector corresponding to the minimum eigenvalue of the clutter covariance matrix is used as the filter weight ($C/N_0 = \infty$). By use of the same filter, I_{cn} is the new IF when white noise is added to the clutter. By taking logarithms of Eq. (22) and approximating for $C/N_0 >> 1$,

$$10 \log I_{cn} \approx 10 \log \frac{C}{N_0} - 10 \log \left[\frac{\frac{C}{N_0}}{I_c} + 1 \right]. \tag{23}$$

If $C/N_0 << I_c$, then $10 \log I_{cn} \approx 10 \log C/N_0$. If $C/N_0 >> I_c$, then $10 \log I_{cn} \approx 10 \log I_c$. Hence C/N_0 is a limiting factor in the improvement factor when white noise is added. Although this relation yields a useful result, it applies to Emerson filters and does not hold for general linear filters. Furthermore, Eq. (22) does not tell us what happens as a function of frequency, only the average gain or improvement factor is presented. A more general result is stated here as a robustness theorem.

Theorem 1 — *Given a vector of filter coefficients let $G_c(\theta)$ represent the gain, Eq. (12) for which R is a $N \times N$ positive definite clutter covariance marix and diagonal elements σ_1^2. Assuming the same filter coefficients, let $G_{cn}(\theta)$ represent the gain for which the covariance matrix is $R' = R + \frac{\sigma_1^2}{\left(\frac{C}{N_0}\right)} I$. If $\Delta(\theta) = 10 \log G_c(\theta) - 10 \log G_{cn}(\theta)$,*

$$10 \log \left[\frac{\frac{C}{N_0}}{1 + \frac{C}{N_0}} \left(1 + \frac{G_{\max}}{N \frac{C}{N_0}} \right) \right] \leq \Delta(\theta) \leq 10 \log \left[\frac{\frac{C}{N_0}}{1 + \frac{C}{N_0}} \left(1 + \frac{G_{\max}}{\frac{C}{N_0}} \right) \right], \quad (24)$$

where $G_{\max} = \max_\theta G_c(\theta)$.

For a proof see Appendix B. If $C/N_0 \geq G_{\max}$, the upper bound indicates $\Delta(\theta)$ is small. So long as $C/N_0 \geq G_{\max}$, any linear filter is robust in the sense that the gain function will not change as a function of C/N_0. If additionally $G_c(\theta)$ is close to the optimal gain in which $C/N_0 = \infty$, then $G_{cn}(\theta)$ will also be close to the optimal gain for every $C/N_0 \geq G_{\max}$. This follows, since the optimal gain for any R' with finite C/N_0 is less then the optimal gain with covariance matrix R. In the case where $C/N_0 \leq G_{\max}$, Eq. (24) yields upper and lower bounds on the degradation and the gain function will suffer. These bounds are plotted in Fig. 15 for $N = 5$, $\sigma_C T = 0.05$, and $C/N_0 = 10, 30, \infty$ dB. The topmost figure is G_c for a single filter designed by the MMLE for $C/N_0 = \infty$. The curves portrayed with $C/N_0 = 10, 30$ dB represent the same filter but with different covariance matrices incorporating the respective white noise. The dotted line curves above and below the $C/N_0 = 10, 30$ dB curves are the upper and lower bounds in Eq. (24). The bounds hold whether or not the interference is Gaussian.

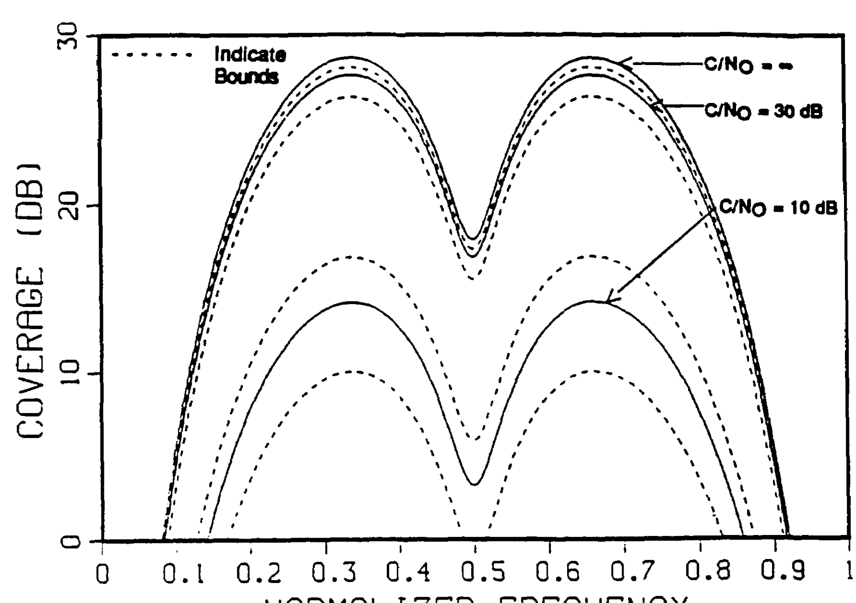

Fig. 15 — Robustness bounds, $N = 5$, $\sigma_C T + 0.05$, $C/N_0 = 10, 30, \infty$ dB

Now we address the question of robustness for the filters considered in the previous sections. For comparison, the filters are all designed assuming $\sigma_C T = 0.05$, $C/N_0 = 90$ dB, with passband $f \in [.2, .8]$ and uniform weighting function in the passband. The dimensionality will be fixed at $N = 10$. The degradation of each filter relative to the infinite bank of filters for different $\sigma_C T$, C/N_0, and f, will be shown.

Figure 16 is a robustness curve of a bank of ten CGLRT filters designed as discussed in Section 6 at $f = 0.5$. Each point of the solid lines represents an optimal filter with weighting $R^{-1}\mathbf{S}^*$ matched at the particular $\sigma_C T$, C/N_0 and f. The dotted lines are the coverage that is obtained with the fixed bank of CGLRT filters designed with the parameters above. This filter bank is particularly robust for $C/N_0 \leq 80$ dB and the large range of $\sigma_C T$ shown. On the other hand, when $f = .2$, Fig. 17, the results are rather poor relative to what can theoretically be achieved. In fact, there are regions where the coverage could increase by 30 dB or more by better processing. Nearly the same coverage is seen in Figs. 18 and 19 for $f = 0.5$, $f = 0.2$ respectively when the design criterion is Andrews' extension of Emerson's filters. Only four filters are used in the MMLE design as shown in Fig. 20 for $f = 0.5$. The filters were designed by partitioning the passband as $f \in [.2, .25]$, $f \in [.25, .5]$, $f \in [.5, .75]$, $f \in [.75, .8]$. The filters are particularly robust at $f = 0.5$ for the parameters shown. In Fig. 21 the coverage is shown for $f = 0.2$. As with the previous cases, the coverage is poor compared with the optimal gain.

The class of filters known as Chebychev is designed somewhat independent of $\sigma_C T, C/N_0$, and f. The Chebychev filters can be characterized by their sidelobe levels and dimensionality [17]. In Fig. 22 the coverage is shown for ten Chebychev filters, equally spaced, Eq. (20), at $f = 0.5$ with 90 dB sidelobe levels. The coverage is mostly identical to the coverage obtained with the previous filters. When $f = 0.2$ as in Fig. 23, the coverage is much better than that obtained in the previous cases. Unfortunately it still lacks the optimal coverage by at most 7 dB at $C/N_0 = 20$ dB, 8 dB at $C/N_0 = 50$ dB, and 23 dB at $C/N_0 = 80$ dB.

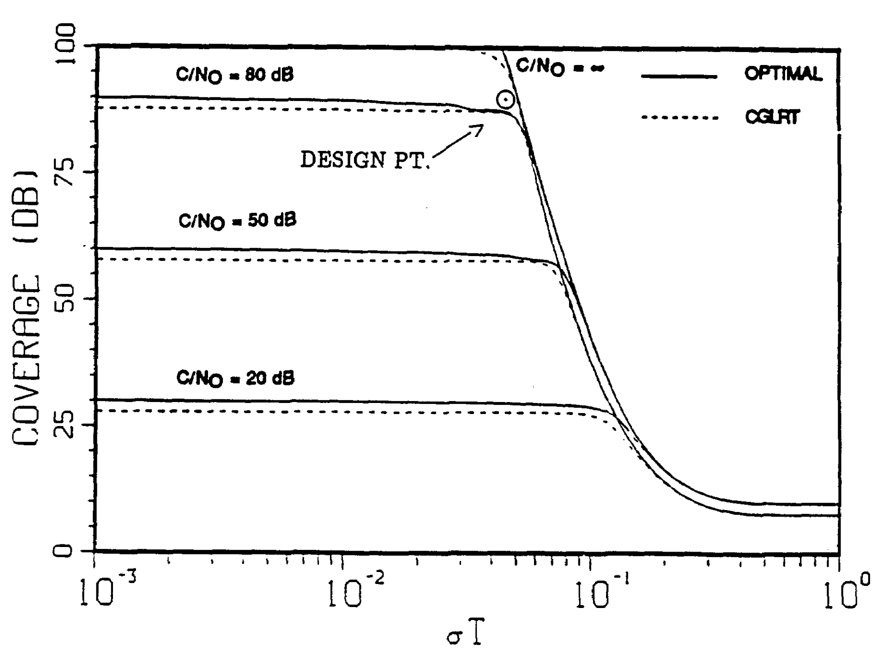

Fig. 16 — Coverage CGLRT $f = 0.5$, $N = 10$ $\sigma_C T = 0.05$, $C/N_0 = 20, 50, 80, \infty$ dB

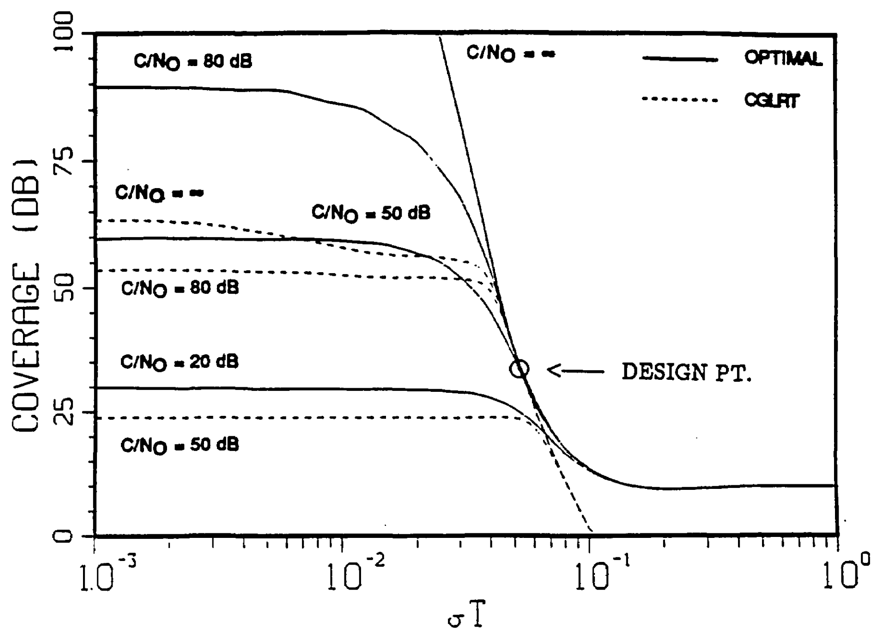

Fig. 17 — Coverage CGLRT $f = 0.2$, $N = 10$ $\sigma_C T = 0.05$, $C/N_0 = 20, 50, 80, \infty$ dB

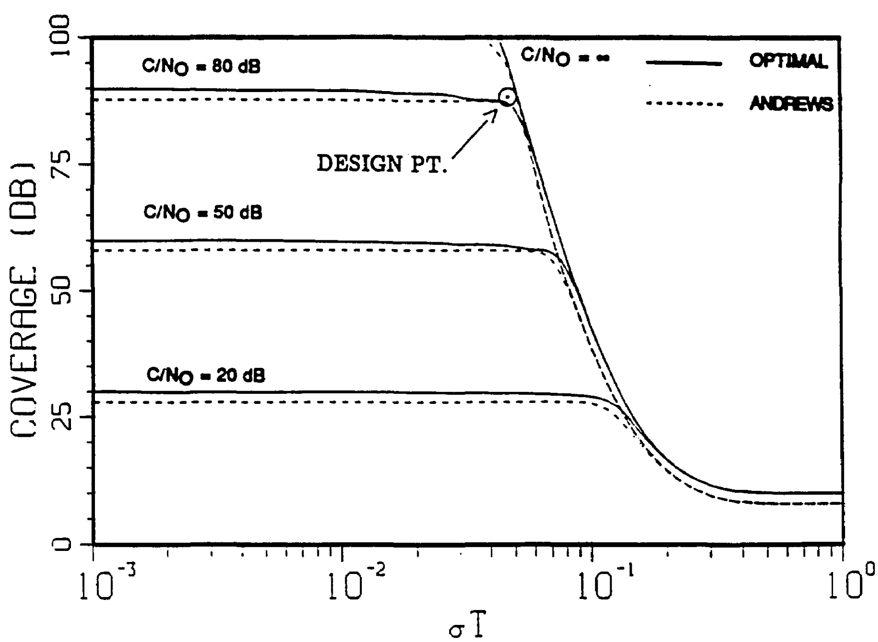

Fig. 18 — Coverage Andrews $f = 0.5$, $N = 10$ $\sigma_C T = 0.05$, $C/N_0 = 20, 50, 80, \infty$ dB

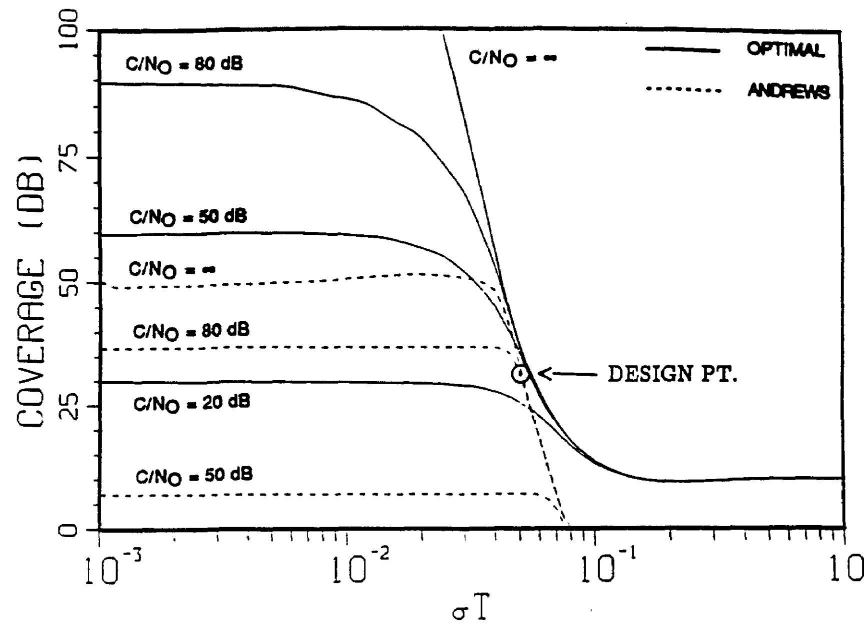

Fig. 19 — Coverage Andrews $f = 0.2$, $N = 10$ $\sigma_C T = 0.05$, $C/N_0 = 20, 50, 80, \infty$ dB

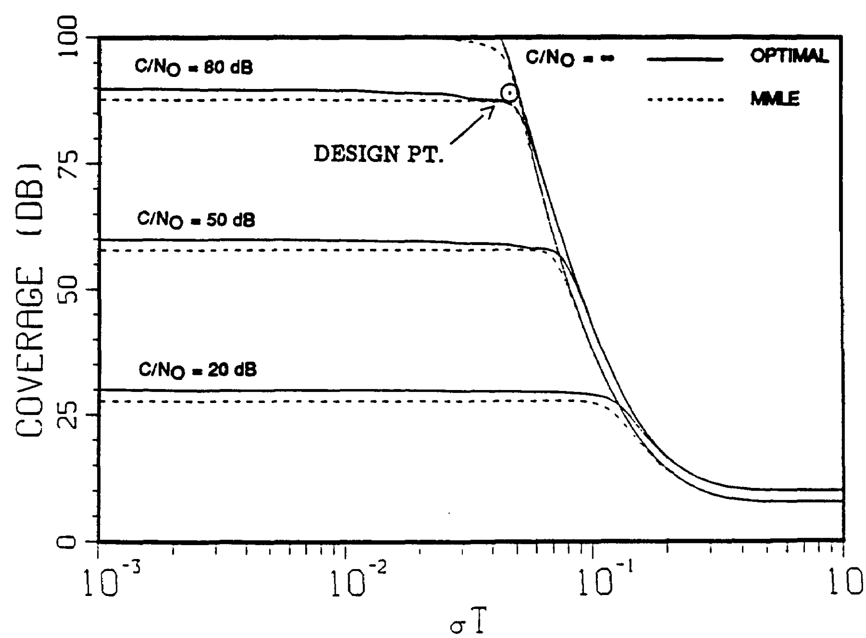

Fig. 20 — Coverage MMLE $f = 0.5$, $N = 10$ $\sigma_C T = 0.05$, $C/N_0 = 20, 50, 80, \infty$ dB

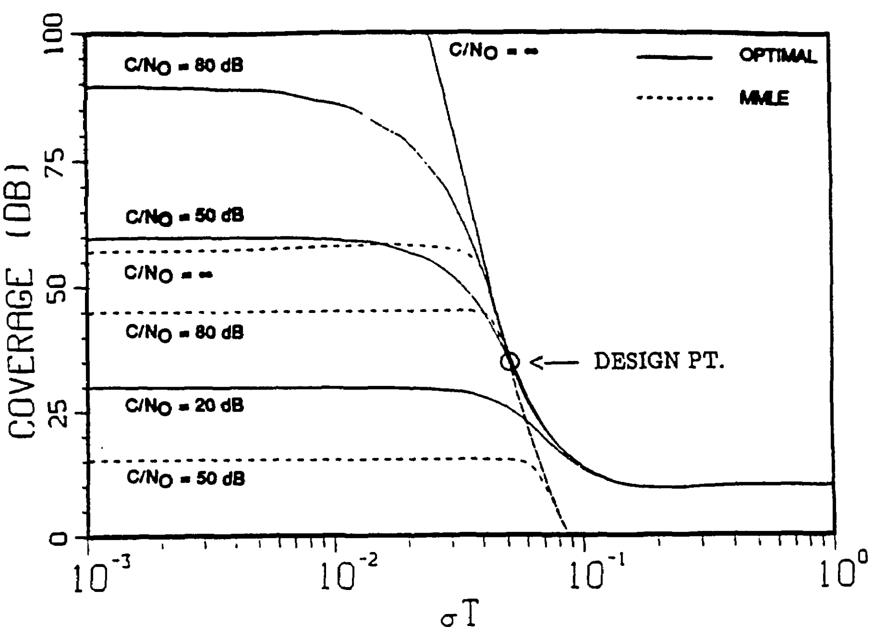

Fig. 21 — Coverage MMLE $f = 0.2$, $N = 10$ $\sigma_C T = 0.05$,
$C/N_0 = 20, 50, 80, \infty$ dB

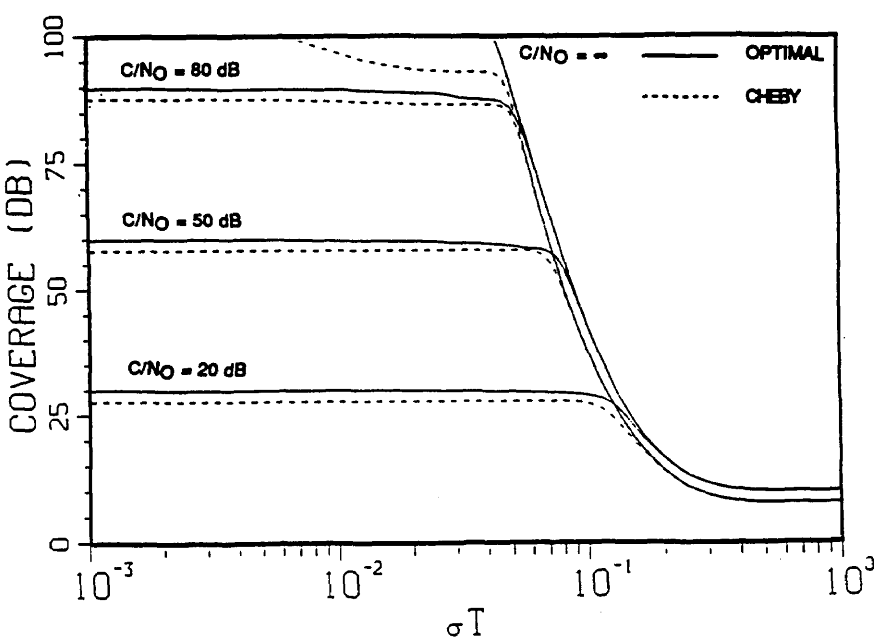

Fig. 22 — Coverage Cheby $f = 0.5$, $N = 10$
$C/N_0 = 20, 50, 80, \infty$ dB

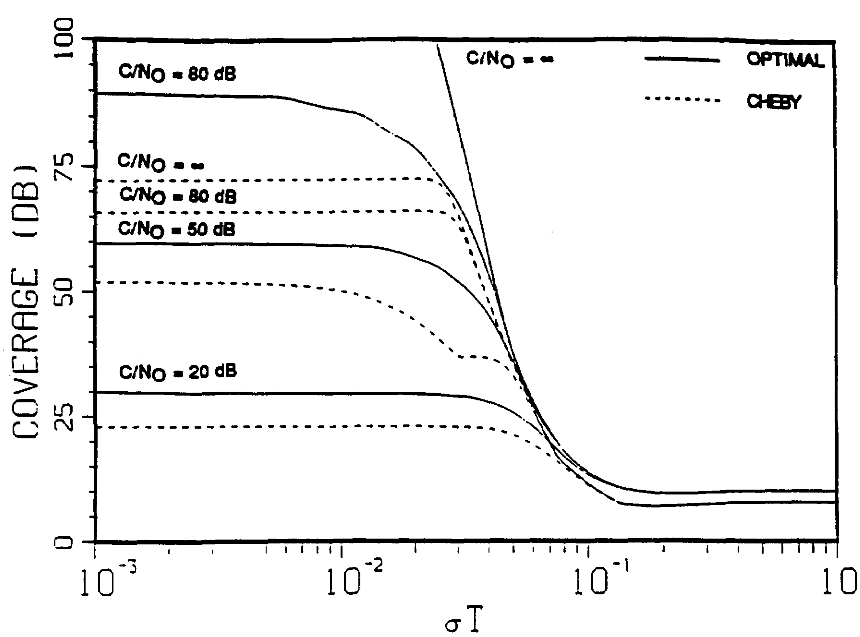

Fig. 23 — Coverage Cheby $f = 0.2$, $N = 10$
$C/N_0 = 20, 50, 80, \infty$ dB

The problem of obtaining filters with frequency coverage nearly optimal for a particular operating point of C/N_0 and $\sigma_C T$ has been solved. When C/N_0 and $\sigma_C T$ are allowed to vary, such filters have been shown to degrade relative to the optimal coverage. This is further examined in the next section.

8. CONCLUSIONS

We have presented a design method that will produce filters capable of detecting targets of unknown velocity. A design criterion is specified by the radar designer in terms of gain relative to the optimal gain as a function of frequency. The method requires partitioning of the passband and comparison of the coverage of the corresponding MMLE filters with the design criteria. Enough filters are added as described in Section 6 until the coverage meets the design criteria. The iterative partitioning is done in an attempt to minimize the number of filters needed. The degree of the closeness was explained in Section 3. The iterative partitioning will very likely produce nonuniform length intervals.

These filters were designed under the premise of known values of $\sigma_C T$ and C/N_0. The case, when these values are not known precisely, is discussed in Section 7. General bounds were presented for a given linear filter when the clutter to thermal noise ratio is allowed to vary. The performance of filters such as the MMLE, which require precise knowledge of the covariance matrix, was determined. This was addressed by introducting a class of covariance matrices $\mathscr{R}$ in which $\sigma_C T$ and C/N_0 are allowed to vary. The actual design of filters that are robust within the covariance matrix class $\mathscr{R}$ is a remaining problem. An important step in designing such filters is to define a criterion appropriate to this problem. Obviously the usefulness of the maximin formulations,

$$\max_{\mathbf{C}} \min_{r \in \mathscr{R}} SIR$$

where the maximization is over the filter coefficients **C**, and MMLE needs to be determined. Additionally alternative criteria [15] should be investigated.

REFERENCES

1. L.E. Brennan, I.S. Reed, and W. Sollfrey, "A Comparison of Average-Likelihood and Maximum-Likelihood Ratio Tests for Detecting Radar Targets of Unknown Doppler Frequency," *IEEE Trans. Inf. Theory* **IT-14** (1), 104-110 (1969).

2. I. Selin, *Detection Theory,* Princeton University Press (Princeton, N.J., 1965).

3. H.V Poor, *An Introduction to Signal Detection and Estimation*, (Springer-Verlag, 1988).

4. R.C. Emerson, "Some Pulse Doppler and MTI techniques," Rand Rept. R-274, Mar. 1954.

5. J. Capon, "Optimum Weighting Functions for the Detection of Sampled Signals in Noise," *IEEE Trans. Inf. Theory* **IT-10** (2), 152-159 (1964).

6. J.K. Hsiao, "On the Optimization of MTI Clutter Rejection," *IEEE Trans. Aerospace and Electron. Syst.* **AES-10** (5), 622-629 (1974).

7. G.A. Andrews, Jr., "Optimization of Radar Doppler Filters to Maximize Moving Target Detection," *NAECON '74 Record,* pp. 279-283.

8. D.C. Schleher and D. Schulkind, "Optimization of Nonrecursive MTI," IEE Int. Radar Conf., London, pp. 182-185 (1977).

9. L.E. Brennan and I.E. Reed, "Optimum Processing of Unequally Spaced Radar Pulse Trains for Clutter Rejection," *IEEE Trans. Aerospace Electron. Syst.* **AES-4** (3), 474-477 (1968).

10. Y. Bresler and A. Macovski, "Exact Maximum Likelihood Parameter Estimation of Superimposed Exponential Signals in Noise", *IEEE Trans. Acoust., Speech, Signal Pro.* **ASSP-34** (5), 1081-1089 (1986).

11. D.V. Compernolle, " Spectral Estimation using a Log-Distance Error Criterion applied to Speech Recognition," *Int. Conf. Acoust., Speech, Signal Proc.* **S6.2**, 258-261 (1989).

12. S.A. Kassam and H.V. Poor, "Robust Techniques for Signal Processing," *Proc. IEEE,* Mar. 1985, pp. 433-481.

13. H.V. Poor, "Robust Matched Filters," *IEEE Trans. Inf. Theory* **IT-29**, 677-687 (1983).

14. L.E. Brennan and I.E. Reed. "Theory of Adaptive Radar," *IEEE Trans. Aerospace Electron. Syst. AES-9* (2), 237-252 (1973).

15. F.F. Kretschmer, "MTI Target Visibility," *IEEE Trans. on Aerospace Electron. Syst.* **AES-22** (2), 216-218 (1986).

16. E. Fong, A. Walker, and W.G. Bath, "Moving Target Indication in the Presense of Radio Frequency Interference," Proc. IEEE Int. Radar Conf., 1985, pp. 292-296.

17. J.K. Hsiao, "FFT Doppler Filter Performance Computations," NRL Report 2744, Mar. 1974.

Appendix A

IMPROVEMENT FACTOR

The criteria of maximizing the IF (improvement factor), referred to as Emerson's criterion, is now examined. The primary consideration is to determine when IF is a good indicator of filter performance. To better understand the criteria, the covariance matrix,

$$R = \sum_i \lambda_i \frac{\mathbf{S}_i \mathbf{S}_i^\dagger}{N}, \quad \mathbf{S} = \begin{bmatrix} 1 \\ e^{j\frac{2\pi(i-1)}{N}} \\ e^{j\frac{4\pi(i-1)}{N}} \\ \cdot \\ \cdot \\ \cdot \\ e^{j\frac{2\pi(N-1)(i-1)}{N}} \end{bmatrix}$$

where $i = 1, \ldots, N$ will be applied. The set $\{\lambda_1, \ldots, \lambda_N\}$ contains the eigenvalues of the matrix. The eigenvalues are assumed distinct and the matrix is Toeplitz. We set the eigenvalues to $\lambda_1 = 1.$, $\lambda_2 = 0.00001$, $\lambda_3 = 0.0000099$, $\lambda_4 = 0.00001$, $\lambda_5 = 0.00001$. Note that when Gaussian interference associated with this covariance matrix is passed through a linear filter, the IF is not $1/\lambda_{\min}$, where $\lambda_{\min}$ is the minimum eigenvalue. This is because the diagonal elements of R are not necessarily equal to one. Assuming the same hypotheses as in Eq. (1) IF $= \dfrac{1}{\gamma}\dfrac{1}{\gamma_{\min}}$, where γ is the input SIR. The solid curve in Fig. 24 is the gain Eq. (12) of the infinite filter bank, Eq. (6). Also shown are curves obtained through the MMLE and Emerson's criteria. Emerson's criterion yields a filter with rather poor gain response. This is because the eigenvector corresponding to the minimum eigenvalue must be $\mathbf{S}_i$ for some i. Every eigenvector is orthogonal to every other eigenvector. Hence, the numerator of Eq. (12) is zero whenever the conjugate of $\mathbf{S}$ is orthogonal to the eigenvector. This happens $N - 1$ times in our example. Essentially $\mathbf{C}^T\mathbf{S}$ where $\theta \in [0, 2\pi]$ represents the conjugate of the discrete time Fourier transform [A1] of the coefficient vector $\mathbf{C}$. Consider the z transform $\Sigma_j c_j z^{-j}$, where z is a complex variable. This equation has $N - 1$ roots z_i, $i = 1, 2, \ldots, N - 1$. If a root is on the unit circle in the z plane $z_i = e^{j\beta}$, the gain will be forced to zero at the frequency $\theta = -\beta$. Robinson [A2] showed that all the roots of the eigenvector corresponding to the minimum eigenvalue are on the unit circle when the covariance matrix is Toeplitz. This may be useful if the roots are close to $\theta = 0$ since it will notch the clutter. However, in cases such as in Fig. A1, much of the remaining Doppler space is also notched.

Emerson's criterion consists of finding the normalized filter coefficients that minimize the interference output power. Equivalently the average gain or IF can be maximized. In the derivation of Emerson's filter, we assume that the Doppler speed is uniformly distributed. Also, no weighting or cost function of frequency exists, i.e., a uniform cost is assumed. The coarse design goal is to produce a filter that delivers favorable gain coverage under these circumstances. One measure for judging the effectiveness of such a criterion is to determine how the IF relates to the design goal. Since the IF is to be maximized, does a higher IF, in fact, indicate a better design? Figure A1 shows that the MMLE yields a filter design uniformly close to the optimal gain. It outperforms Emerson's filter in 80% of the Doppler space and does not suffer the numerous holes in the coverage. Yet, for the MMLE design, IF = 42.997 dB, and for Emerson's design IF = 43.000 dB. In this example, the IF does not effectively relate to the design goal. To further illustrate this, examine Fig. A2 for $N = 5$, $\sigma_c T = 0.05$, $c/N_0 = 50$ dB. The solid curve was obtained from a single filter design by MMLE with passband $f \in [.2,.8]$ and uniform weighting. It was determined the IF = 43.76 dB. Also shown in the figure by the dotted lines is a rectangular window. This window has no gain in $f \in [0,.4]$, $f \in [.6,1]$ and performs slightly better then the MMLE in $f \in [.4,.6]$. Obviously this would make a rather poor filter compared with the MMLE. The IF is 43.82 dB, which is slightly higher then the MMLE IF. Although the actual eigenvector solution performs much better, and an actual filter that yields the dotted gain curve may not even exist, the example does help to show that IF is not necessarily a good indicator of the design goal.

In summary, Emerson's criterion works well when applied to problems where the covariance matrix is such that the above problems do not exist. When these problems arise, as in Fig. A1 or in the case of extraneous interference as discussed in Section 5, an alternative criterion should be applied.

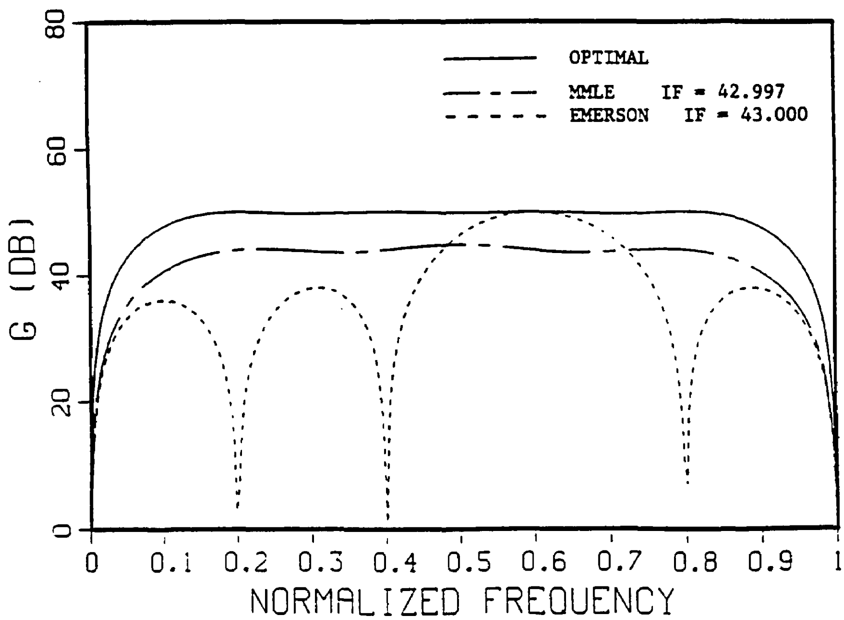

Fig. A1 — Interference matrix (17), $N = 5$

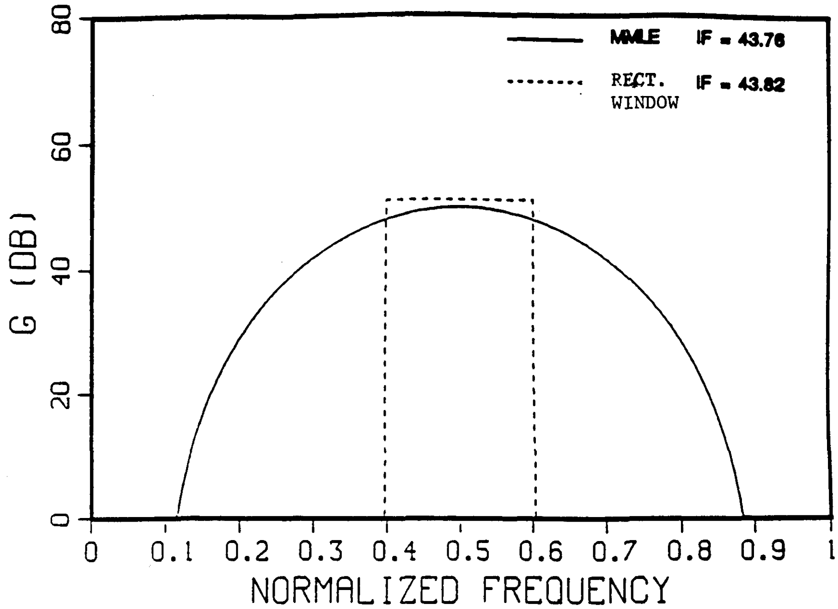

Fig. A2 — Improvement factors

REFERENCES

A1. A.V. Oppenheim and R.W. Schafer, *Digital Signal Processing,* (Prentice-Hall, Inc., New Jersey, 1975).

A2. E.A. Robinson, "On Optimum Weighting Functions for the Detection of Sampled Signals in Noise," *IEEE Trans. Inf. Theory* **IT-11**, 452-453 (1965).

Appendix B

PROOF OF ROBUSTNESS THEOREM

Define the vectors $\mathbf{S}(\theta)$, $\mathbf{S}_i, S_i'$ by

$$\mathbf{S}(\theta) = \begin{bmatrix} 1 \\ e^{j\theta} \\ \cdot \\ \cdot \\ \cdot \\ e^{j(N-1)\theta} \end{bmatrix} \quad \mathbf{S}_i = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 \\ e^{\frac{j2\pi i}{N}} \\ \cdot \\ \cdot \\ \cdot \\ e^{\frac{j2\pi i(N-1)}{N}} \end{bmatrix} \quad \mathbf{S}_i' = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 \\ e^{-\frac{j2\pi i}{N}} \\ \cdot \\ \cdot \\ \cdot \\ e^{\frac{-j2\pi i(N-1)}{N}} \end{bmatrix}$$

where $i = 1, 2, \ldots, N$ and $\theta \in [0,2\pi]$. It can be shown that the sets $\{\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_N\}$ and $\{\mathbf{S}_1', \mathbf{S}_2' \ldots, \mathbf{S}_N'\}$ are orthonormal bases for complex N-space C^N. This additional set is introduced for convenience in the proof of the theorem.

Lemma 1 — *Let $w \in C^N$ such that $w^T w^* = 1$. Then $1 \le \max_\theta | w^T \mathbf{S}(\theta) |^2 \le N$. Equality on the right is achieved if $w \,\alpha\, \mathbf{S}(-\theta_{\max})$, where $\theta_{\max}$ is any θ, where $\max_\theta | w^T \mathbf{S}(\theta) |^2$ is achieved. Equality on the left is achieved if $w^T \mathbf{S}_1 = w^T \mathbf{S}_2 = \ldots = w^T \mathbf{S}_N = 1/\sqrt{N}$.*

Proof: The inequality $\max_\theta | w^T \mathbf{S}(\theta) |^2 \le N$ is argued first. Since $\mathbf{S}_1, \ldots, \mathbf{S}_N$ is a basis for C^N, we can write

$$w = \sum a_i \mathbf{S}_i', \; \sum | a_i |^2 = 1$$

$$\mathbf{S}(\theta) = \sum b_i(\theta) \mathbf{S}_i, \; \sum | b_i(\theta) |^2 = N. \tag{B1}$$

Now

$$| w^T \mathbf{S}(\theta) |^2 = \left| \left[\sum_i a_i \mathbf{S}_i' \right]^T \left[\sum_j b_j \mathbf{S}_j \right] \right|^2$$

$$= \left| \sum_i a_i b_i \right|^2 ,$$

but by the Cauchy-Schwartz inequality,

$$\leq \left[\sum_i |a_i|^2\right] \left[\sum_j |b_j|^2\right]$$

$$= N$$

with equality for a given θ if and only if $\omega \alpha S(-\theta)$. Since $|w^T S(\theta)|^2$ is a continuous function over a closed interval, the maximum exists and is realized at some $\theta_{max} \in [0,2\pi]$. Consequently the preceding inequality implies $\max_\theta |w^T S(\theta)|^2 \leq N$, with equality if $\omega \alpha S(-\theta_{max})$.

Next it is proven that $1 \leq \max_\theta |w^T S(\theta)|^2$. Since $\{S_1', S_2' \ldots , S_N'\}$ is an orthonormal basis, $w = \sum_i a_i S_i'$ and therefore $a_i = w^T S_i$ for $i \in \{1,2, \ldots , N\}$. Observe that $S(\theta_i) = \sqrt{N} S_i$ for $\theta_i = 2\pi i/N$, from which it follows that $|w^T S(\theta_i)| = \sqrt{N} a_i$ and

$$\max_0 |w^T S(\theta)| \geq \sqrt{N} |a_i|$$

for all i. Hence

$$\max_\theta |w^T S(\theta)| \geq \max \{\sqrt{N} |a_1|, \sqrt{N} |a_2|, \ldots , \sqrt{N} |a_N|\}.$$

To finish the argument consider two cases: (a) $|a_i| \geq 1/\sqrt{N}$ for all $i \in \{1, 2, \ldots , N\}$ and (b) $|a_i| < 1/\sqrt{N}$ for some $i \in \{1, 2, \ldots , N\}$. In case (a) the result follows trivially. In case (b) $|a_i| \geq 1/\sqrt{N}$ for at least one of the subscripts $i \in \{1, 2 , \ldots , N\}$; otherwise $\sum_i |a_i|^2 < 1$, which is a contradiction. Therefore the maximum on the right-hand side is at least as large as 1. Hence

$$\max_\theta |w^T S(\theta)| \geq 1$$

and consequently

$$\max_\theta |w^T S(\theta)|^2 \geq 1.$$

Finally equality on the left is argued if $a_1 = a_2 = \ldots = a_N = 1/\sqrt{N}$:

$$\max_\theta |w^T S(\theta)|^2 = \max_\theta \left| \left(\sum a_i (S_i')^T\right) S(\theta)\right|^2$$

$$= \max \left| \frac{1}{\sqrt{N}} \left[\sum_i (S_i')^T\right] \left[\sum_j b_j(\theta) S_j\right] \right|^2$$

where b_j are as in (a). By orthonormality this becomes,

$$= \frac{1}{N} \max_{\theta} \left| \sum b_j(\theta) \right|^2.$$

Applying the Cauchy-Schwartz inequality,

$$\leq \frac{1}{N} \max_{\theta} \sum \left| b_j(\theta) \right|^2 = 1.$$

The result follows since $\max_\theta \left| w^T \mathbf{S}(\theta) \right|^2 \geq 1$.

Proof of Robustness Theorem:

Using the definitions of Theorem 1 and without loss of generality, we assume the filter vector w satisfies $w^T w^* = 1$. Then,

$$\begin{aligned}
\Delta(\theta) &= 10 \log G_c(\theta) - 10 \log G_{cn}(\theta) \\
&= 10 \log \left[\frac{1}{\gamma_c} \frac{\left| w^T \mathbf{S}(\theta) \right|^2}{w \dagger R w} \right] - 10 \log \left[\frac{1}{\gamma_{cn}} \frac{\left| w^T \mathbf{S}(\theta) \right|^2}{w \dagger R' w} \right] \\
&= 10 \log \left[\frac{\gamma_{cn}}{\gamma_c} \frac{w \dagger R' w}{w \dagger R w} \right] \\
&= 10 \log \frac{\gamma_{cn}}{\gamma_c} \left[1 + \frac{\sigma_1^2}{\dfrac{C}{N_0}} \frac{1}{w \dagger R w} \right],
\end{aligned} \tag{B2}$$

where the input SIRS corresponding to the covariance matrices R and R' are $\gamma_c = \dfrac{\mathbf{S}^T \mathbf{S}^*}{N \sigma_1^2} = \dfrac{1}{\sigma^2}$ and $\gamma_{cn} = \dfrac{\mathbf{S}^T \mathbf{S}^*}{N \sigma_1^2 \left[1 + \dfrac{1}{\dfrac{C}{N_0}} \right]} = \dfrac{1}{\sigma_1^2 \left[1 + \dfrac{1}{\dfrac{C}{N_0}} \right]}$, respectively. Note that G is a continuous function of θ over $[0, 2\pi]$. Thus the maximum of G is obtained for some θ in the interval denoted by $\theta_{\max}$, and

$$\begin{aligned}
G_{\max} &= \max_{\theta} G_c(\theta) \\
&= G(\theta_{\max}) \\
&= \frac{1}{\gamma_c} \frac{\left| w^T \mathbf{S}(\theta_{\max}) \right|^2}{w \dagger R w}.
\end{aligned}$$

This can be rewritten as

$$\max_{\theta} |w^T \mathbf{S}(\theta)|^2 = |w^T \mathbf{S}(\theta_{\max})|^2 = \gamma_c G_{\max} w^{\dagger} R w.$$

Applying the lemma yields the inequality

$$1 \le \gamma_c G_{\max} w^{\dagger} R w \le N,$$

or equivalently,

$$\frac{\gamma_c G_{\max}}{N} \le \frac{1}{w^{\dagger} R w} \le \gamma_c G_{\max},$$

since R is positive definite and $\gamma_c > 0$. Combining this with Eq. (B2),

$$10 \log \left[\frac{\gamma_{cn}}{\gamma_c} \left(1 + \frac{\sigma_1^2 \gamma_c}{N \dfrac{C}{N_0}} G_{\max} \right) \right] \le \Delta(\theta) \le 10 \log \left[\frac{\gamma_{cn}}{\gamma_c} \left(1 + \frac{\sigma_1^2 \gamma_c}{\dfrac{C}{N_0}} G_{\max} \right) \right]$$

or since $\dfrac{\gamma_{cn}}{\gamma_c} = \dfrac{\dfrac{C}{N_0}}{\dfrac{C}{N_0} + 1}$,

$$10 \log \left[\frac{\dfrac{C}{N_0}}{1 + \dfrac{C}{N_0}} \left(1 + \frac{1}{N \dfrac{C}{N_0}} G_{\max} \right) \right] \le \Delta(\theta) \le 10 \log \left[\frac{\dfrac{C}{N_0}}{1 + \dfrac{C}{N_0}} \left(1 + \frac{1}{\dfrac{C}{N_0}} G_{\max} \right) \right]$$

Appendix C

COMPUTATION OF MMLE FILTERS

The computation of

$$\min_{C}\left[\mathbf{C}^{\dagger} R\mathbf{C} \max_{\theta\in\Lambda_A} \frac{\mathbf{S}^T R^{-1}\,\mathbf{S}^*}{|\mathbf{C}^T\mathbf{S}|^2}\right]$$

and the corresponding $\mathbf{C}$ is important since, to the authors knowledge, a general solution has not been found. However, a number of numerical methods are available to compute this. Two algorithms are briefly mentioned here, the first based on simulated annealing, the second on a deterministic method called the directed reduction method.

The simulated annealing algorithm [C1] involves an energy function

$$E\{i\} = \min_{C\{i\}}\left[\mathbf{C}\{i\}^{\dagger} R\mathbf{C}\{i\} \max_{\theta\in\Lambda_A} \frac{\mathbf{S}^T R^{-1}\mathbf{S}^*}{|\mathbf{C}\{i\}^T\mathbf{S}|^2}\right],$$

and temperature,

$$T\{i\} = \frac{T\{0\}}{\log(1+i)},\ i = 1,2,\ldots,$$

where $E\{i\}$ is the energy, $\mathbf{C}\{i\}$ is the filter coefficient vector, and $T\{i\}$ is the temperature at time i. The inner maximization over θ can be approximated by uniformly partitioning $\theta \in [0, 2\pi]$ and discretizing θ. Denoting this discretized set as $\Lambda_{A'} = \{\theta_1, \theta_2, \ldots, \theta_S\}$ the function becomes

$$E\{1\} = \min_{C\{i\}}\left[\mathbf{C}\{i\}^{\dagger} R\mathbf{C}\{i\} \max_{\theta\in\Lambda_{A'}} \frac{\mathbf{S}^T R^{-1}\mathbf{S}^*}{|\mathbf{C}\{i\}^T\mathbf{S}|^2}\right]. \qquad \text{(C1)}$$

For the examples shown in this report $\mathbf{S}$ was between seven and ten. This value, however, may need to be increased if the covariance matrix is complicated or N is large. New coefficients $\mathbf{C}'$ are computed as,

$$\mathbf{C}' = \mathbf{C}\{i\} + \mathbf{V}\{i\}$$

where $V\{i\}$ is a vector of values at time i where each component, at each time, is chosen randomly by a zero mean IID Gaussian distribution with variance proportional to $T\{i\}$. The change in energy is computed,

$$\Delta E = E\{i+1\} - E\{i\}.$$

If $\Delta E < 0$ the new coefficients are accepted, $\mathbf{C}\{i+1\} = \frac{\mathbf{C}'}{(\mathbf{C}')^T(\mathbf{C}')^*}$. If $\Delta E \geq 0$ the coefficients are accepted with probability p,

$$p = e^{\frac{-\Delta E}{T\{i\}}};$$

i.e., a random number r is generated uniformly in the interval [0, 1] and if $r < p$ the coefficient vector is $\mathbf{C}\{i+1\} = \frac{\mathbf{C}'}{(\mathbf{C}')^T(\mathbf{C}')^*}$, otherwise $\mathbf{C}\{i+1\} = \mathbf{C}\{i\}$, $E\{i+1\} = E\{i\}$. This is continued until the temperature is reduced to the point where the energy reaches steady state. Simulated annealing algorithms such as this are fairly effective and easily programmed. Most filters designed in this report took between 10 and 60 min of running time to approximately reach steady state on a Sun 3/60 computer. The running time is very sensitive to the initial conditions so it is best to make a good guess before beginning. Typically filters of the class $R^{-1}\mathbf{S}^*$ with the particular $\theta \in \Lambda_{A'}$ that minimize Eq. (C1) are useful for this purpose.

The faster directed reduction method was devised as described. For each of the $\theta_i \in \Lambda_{A'}$ optimal $R^{-1}\mathbf{S}^*$ filter coefficients matched at that frequency are predetermined. Denote these coefficients as $\{\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_S\}$ that are optimal and correspond to $\{\theta_1, \theta_2, \ldots, \theta_S\}$ respectively. At each step i, the energy is computed in Eq. (C1). Suppose the inner maximum occurs at $\theta = \theta_l, l \in \{1,2,\ldots,S\}$, the new coefficients are then determined as

$$\mathbf{C}' = \mathbf{C}\{i\} + k\mathbf{C}_l$$

$$\mathbf{C}\{i+1\} = \frac{\mathbf{C}'}{(\mathbf{C}')^T(\mathbf{C}')^*},$$

where $\mathbf{C}_l$ is the optimal weighting corresponding to θ_l. Note that the overall energy Eq. (C1) may increase or decrease because of this change. It should be monitored to see if steady state is reached. This algorithm is not particularly sensitive to the initial conditions relating to the time to reach steady state, assuming the algorithm converges. It is, on the other hand, less stable than the simulated annealing algorithm, and the initial conditions are important for convergence to occur. The parameter k controls the speed of convergence and is related to the algorithm's stability. If the value is too large, the algorithm may be unstable. For the applications in this report $0.01 \geq k \geq 0.0001$. The filters designed for this report took between 1 and 15 min of running time to approximately reach steady state on a Sun 3/60 computer.

REFERENCE

C1. N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller. "Equation of State Calculations by Fast Computing Machines," *J. Chem. Phys.* **21**, 1087-1092 (1953).